PENNIE & EDMONDS LLP
COUNSELLORS AT LAW
1155 Avenue of the Americas
New York, New York 10036
(212) 790-9090

ATTORNEY DOCKET NO. 7853-114 Date March 4, 1998

Assistant Commissioner for Patents
Box PATENT APPLICATION
Washington, D.C. 20231

Sir:

The following utility patent application is enclosed for filing:

Applicant(s): Dean A. FALB Executed on: *unexecuted*

Title of Invention: COMPOSITIONS AND METHODS FOR THE TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE

PATENT APPLICATION FEE VALUE

TYPE	NO. FILED	LESS	EXTRA	EXTRA RATE	FEE
Total Claims	130	-20	110	$22.00 each	2,420.00
Independent	17	-3	14	$82.00 each	1,148.00
			Minimum Fee		790.00
			Multiple Dependency Fee If Applicable ($270.00)		270.00
			Total		4,628.00
			50% Reduction for Independent Inventor, Nonprofit Organization or Small Business Concern (a verified statement as to the applicant's status is attached)		-
				Total Filing Fee	$ 4,628.00

☐ Priority of application no. filed on in U.S.A. is claimed under 35 U.S.C. § 119.
☐ The certified copy of the priority application has been filed in application no. filed

Please charge the required fee to Pennie & Edmonds LLP Deposit Account No. 16-1150. A copy of this sheet is enclosed.

Respectfully submitted,

Jonathan L. Klein 41,119
Jonathan L. Klein (Reg. No.)

For: Laura A. Coruzzi
Reg. No.: 30,742

PENNIE & EDMONDS LLP

Enclosure

This form is not for use with continuation, divisional, re-issue, design or plant patent applications.

A

Please type a plus (+) inside this box → ⊞

JC544 U.S. PTO 03/04/98

PTO/SB/29 (12/97)
Approved for use through 09/30/00. OMB 0651-0032
Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

UTILITY PATENT APPLICATION TRANSMITTAL

(Only for new nonprovisional applications under 37 CFR 1.53(b))

Attorney Docket No.	7853-114	*Total Pages*	209
First Named Inventor or Application Identifier	Dean A. FALB		
Express Mail Label No.			

APPLICATION ELEMENTS
See MPEP chapter 600 concerning utility patent application contents.

ADDRESS TO: **Assistant Commissioner for Patents**
Box Patent Application
Washington, DC 20231

1. ☒ Fee Transmittal Form
   Submit an original, and a duplicate for fee processing)
2. ☒ Specification [*Total Pages* 180 (plus 5 pages table of contents)]
   (preferred arrangement set forth below)
   - -Descriptive title of the Invention
   - -Cross Reference to Related Applications
   - -Statement Regarding Fed sponsored R&D
   - -Reference to Microfiche Appendix
   - -Background of the Invention
   - -Brief Summary of the Invention
   - -Brief Description of the Drawings *(if filed)*
   - -Claim(s)
   - -Abstract of the Disclosure
3. ☒ Drawing(s) *(35 USC 113)* [*Total Sheets 21*]
4. ☒ Oath or Declaration [*Total Sheets 2*]
   a. ☒ Unexecuted
   b. ☐ Copy from a prior application (37 CFR 1.63(d))
   (for continuation/divisional with Box 17 completed)
   [Note Box 5 below]
   i. ☐ DELETION OF INVENTOR(S)
   Signed statement attached deleting inventor(s) named in the prior application, see 37 CFR 1.63(d)(2) and 1.33 (b).
5. ☐ Incorporation By Reference *(useable if Box 4b is checked)*
   The entire disclosure of the prior application, from which a copy of the oath or declaration is supplied under Box 4b, is considered as being part of the disclosure of the accompanying application and is hereby incorporated by reference therein.
6. ☐ Microfiche Computer Program *(Appendix)*
7. ☐ Nucleotide and/or Amino Acid Sequence Submission *(if applicable, all necessary)*
   a. ☐ Computer Readable Copy
   b. ☐ Paper Copy (identical to computer copy)
   c. ☐ Statement verifying identity of above copies

ACCOMPANYING APPLICATION PARTS

8. ☐ Assignment Papers (cover sheet & document(s))
9. ☐ 37 CFR 3.73(b) Statement ☒ Power of Attorney
   (when there is an assignee) *(unexected)*
10. ☐ English Translation Document *(if applicable)*
11. ☐ Information Disclosure Statement (IDS)/PTO-1449 ☐ Copies of IDS Citations
12. ☐ Preliminary Amendment
13. ☒ Return Receipt Postcard (MPEP 503)
    (Should be specifically itemized)
14. ☐ Small Entity ☐ Statement filed in prior application, Statement(s) Status still proper and desired
15. ☐ Certified Copy of Priority Document(s)
    (if foreign priority is claimed)
16. ☐ Other:

17. **If a CONTINUING APPLICATION,** *check appropriate box and supply the requisite information:*
☐ Continuation ☐ Divisional ☒ Continuation-in-part (CIP) of prior application No: 08/870,434, filed June 6, 1997, which is a continuation-in-part of application No. 08/799,910, filed February 13, 1997, which claims the benefit under 35 U.S.C. § 119(e) of provisional application No. 60/011,787, filed February 16, 1996, now abandoned; and this application is a continuation-in-part of application No. 08/599,654, filed February 9, 1996, which is a continuation-in-part of application No. 08/485,573, filed June 7, 1995, which is a continuation-in-part of application No. 08/386,844, filed February 10, 1995.

18. CORRESPONDENCE ADDRESS

☒ Customer Number or Bar Code Label 20582 *(Insert Customer No. or Attach bar code label here)* or ☐ Correspondence address below

NAME					
ADDRESS					
CITY		*STATE*		*ZIP CODE*	
COUNTRY		*TELEPHONE*		*FAX*	

Burden Hour Statement: This form is estimated to take 0.2 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Box Patent Application, Washington, DC 20231.

COMPOSITIONS AND METHODS FOR THE TREATMENT
AND DIAGNOSIS OF CARDIOVASCULAR DISEASE

TABLE OF CONTENTS

Page

COMPOSITIONS AND METHODS FOR THE TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE

This application is a continuation-in-part of co-pending application serial number 08/870,434, filed June 6, 1997, which is a continuation-in-part of co-pending application serial number 08/799,910, filed February 13, 1997, which claims the benefit under 35 U.S.C. § 119(e) of provisional application serial number 60/011,787, filed February 16, 1996, now abandoned; and this application is a continuation-in-part of co-pending application serial number 08/599,654, filed February 9, 1996, which is a continuation-in-part of co-pending application serial number 08/485,573, filed June 7, 1995, which is a continuation-in-part of co-pending application serial number 08/386,844, filed February 10, 1995, each of which is hereby incorporated by reference in its entirety.

1. INTRODUCTION

The present invention relates to methods and compositions for the treatment and diagnosis of cardiovascular disease, including, but not limited to, atherosclerosis, ischemia/reperfusion, hypertension, restenosis, and arterial inflammation. The present invention further relates to screening methods to identify compositions and their therapeutic use for the treatment of fibro-proliferative and oncogenic disorders, including diabetic retinopathy, artherosclerosis, angiogenesis, inflammation, fibrosis, tumor growth and vascularization. Genes which are differentially expressed in cardiovascular or oncogenic disease states, relative to their expression in normal, or non-disease states are identified. Genes are also identified via the ability of their gene products to interact with other gene products involved in cardiovascular or oncogenic disease. The genes identified may be used diagnostically or as targets for therapeutic intervention. In this regard, the present invention provides methods for the identification and

therapeutic use of compounds in the treatment and diagnosis of cardiovascular disease. Additionally, methods are provided for the diagnostic monitoring of patients undergoing clinical evaluation for the treatment of cardiovascular disease, for monitoring the efficacy of compounds in clinical trials, and for identifying subjects who may be predisposed to cardiovascular disease.

2. BACKGROUND OF THE INVENTION

Cardiovascular disease is a major health risk throughout the industrialized world. Atherosclerosis, the most prevalent of cardiovascular diseases, is the principal cause of heart attack, stroke, and gangrene of the extremities, and thereby the principal cause of death in the United States. Atherosclerosis is a complex disease involving many cell types and molecular factors (for a detailed review, see Ross, 1993, Nature 362: 801-809). The process, in normal circumstances a protective response to insults to the endothelium and smooth muscle cells (SMCs) of the wall of the artery, consists of the formation of fibrofatty and fibrous lesions or plaques, preceded and accompanied by inflammation. The advanced lesions of atherosclerosis may occlude the artery concerned, and result from an excessive inflammatory-fibroproliferative response to numerous different forms of insult. For example, shear stresses are thought to be responsible for the frequent occurrence of atherosclerotic plaques in regions of the circulatory system where turbulent blood flow occurs, such as branch points and irregular structures.

The first observable event in the formation of an atherosclerotic plaque occurs when blood-borne monocytes adhere to the vascular endothelial layer and transmigrate through to the sub-endothelial space. Adjacent endothelial cells at the same time produce oxidized low density lipoprotein (LDL). These oxidized LDL's are then taken up in large amounts by the monocytes through scavenger receptors expressed on their surfaces. In contrast to the regulated

pathway by which native LDL (nLDL) is taken up by nLDL specific receptors, the scavenger pathway of uptake is not regulated by the monocytes.

These lipid-filled monocytes are called foam cells, and are the major constituent of the fatty streak. Interactions between foam cells and the endothelial and SMCs which surround them lead to a state of chronic local inflammation which can eventually lead to smooth muscle cell proliferation and migration, and the formation of a fibrous plaque. Such plaques occlude the blood vessel concerned and thus restrict the flow of blood, resulting in ischemia.

Ischemia is a condition characterized by a lack of oxygen supply in tissues of organs due to inadequate perfusion. Such inadequate perfusion can have number of natural causes, including atherosclerotic or restenotic lesions, anemia, or stroke, to name a few. Many medical interventions, such as the interruption of the flow of blood during bypass surgery, for example, also lead to ischemia. In addition to sometimes being caused by diseased cardiovascular tissue, ischemia may sometimes affect cardiovascular tissue, such as in ischemic heart disease. Ischemia may occur in any organ, however, that is suffering a lack of oxygen supply.

The most common cause of ischemia in the heart is atherosclerotic disease of epicardial coronary arteries. By reducing the lumen of these vessels, atherosclerosis causes an absolute decrease in myocardial perfusion in the basal state or limits appropriate increases in perfusion when the demand for flow is augmented. Coronary blood flow can also be limited by arterial thrombi, spasm, and, rarely, coronary emboli, as well as by ostial narrowing due to luetic aortitis. Congenital abnormalities, such as anomalous origin of the left anterior descending coronary artery from the pulmonary artery, may cause myocardial ischemia and infarction in infancy, but this cause is very rare in adults. Myocardial ischemia can also occur if myocardial oxygen demands are abnormally increased, as in severe ventricular

hypertrophy due to hypertension or aortic stenosis. The latter can be present with angina that is indistinguishable from that caused by coronary atherosclerosis. A reduction in the oxygen-carrying capacity of the blood, as in extremely severe anemia or in the presence of carboxy-hemoglobin, is a rare cause of myocardial ischemia. Not infrequently, two or more causes of ischemia will coexist, such as an increase in oxygen demand due to left ventricular hypertrophy and a reduction in oxygen supply secondary to coronary atherosclerosis.

The principal surgical approaches to the treatment of ischemic atherosclerosis are bypass grafting, endarterectomy, and percutaneous translumenal angioplasty (PCTA). The failure rate after these approaches due to restenosis, in which the occlusions recur and often become even worse, is extraordinarily high (30-50%). It appears that much of the restenosis is due to further inflammation, smooth muscle accumulation, and thrombosis.

Very recently, a modified balloon angioplasty approach was used to treat arterial restenosis in pigs by gene therapy (Ohno et al., 1994, Science 265: 781-784). A specialized catheter was used to introduce a recombinant adenovirus carrying the gene encoding thymidine kinase (tk) into the cells at the site of arterial blockage. Subsequently, the pigs were treated with ganciclovir, a nucleoside analog which is converted by tk into a toxic form which kills cells when incorporated into DNA. Treated animals had a 50% to 90% reduction in arterial wall thickening without any observed local or systemic toxicities.

Because of the presumed role of the excessive inflammatory-fibroproliferative response in atherosclerosis and ischemia, a number of researchers have investigated, in the context of arterial injury, the expression of certain factors involved in inflammation, cell recruitment and proliferation. These factors include growth factors, cytokines, and other chemicals, including lipids involved in

cell recruitment and migration, cell proliferation and the control of lipid and protein synthesis.

For example, the expression of PDGF (platelet derived growth factor) or its receptor was studied: in rats during repair of arterial injury (Majesky et al., 1990, J. Cell Biol. 111: 2149); in adherent cultures of human monocyte-derived macrophages treated with oxidized LDL (Malden et al., 1991, J. Biol. Chem. 266: 13901); and in bovine aortic endothelial cells subjected to fluid shear stress (Resnick et al., 1993, Proc. Natl. Acad. Sci. USA 90: 4591-4595). Expression of IGF-I (insulin-like growth factor-I) was studied after balloon deendothelialization of rat aorta (Cercek et al., 1990, Circulation Research 66: 1755-1760).

Other studies have focused on the expression of adhesion-molecules on the surface of activated endothelial cells which mediate monocyte adhesion. These adhesion molecules include intracellular adhesion molecule-1, ICAM-1 (Simmons et al., 1988, Nature, 331: 624-627), ELAM (Bevilacqua et al., 1989, Science 243: 1160-1165; Bevilacqua et al., 1991, Cell 67: 233), and vascular cell adhesion molecule, VCAM-1 (Osborn et al., 1989, Cell 59: 1203-1211); all of these surface molecules are induced transcriptionally in the presence of IL-1. Histological studies reveal that ICAM-1, ELAM and VCAM-1 are expressed on endothelial cells in areas of lesion formation in vivo (Cybulsky et al., 1991, Science 251: 788-791; 1991, Arterioscler. Thromb. 11: 1397a; Poston et al., 1992, Am. J. Pathol. 140: 665-673). VCAM-1 and ICAM-1 were shown to be induced in cultured rabbit arterial endothelium, as well as in cultured human iliac artery endothelial cells by lysophophatidylcholine, a major phospholipid component of atherogenic lipoproteins. (Kume et al., 1992, J. Clin. Invest. 90: 1138-1144). VCAM-I, ICAM-1, and class II major histocompatibility antigens were reported to be induced in response to injury to rabbit aorta (Tanaka, et al., 1993, Circulation 88: 1788-1803).

Recently, cytomegalovirus (CMV) has been implicated in restenosis as well as atherosclerosis in general (Speir, et al., 1994, Science 265: 391-394). It was observed that the CMV protein IE84 apparently predisposes smooth muscle cells to increased growth at the site of restenosis by combining with and inactivating p53 protein, which is known to suppress tumors in its active form.

The foregoing studies are aimed at defining the role of particular gene products presumed to be involved in the excessive inflammatory-fibroproliferative response leading to atherosclerotic plaque formation. However, such approaches cannot identify the full panoply of gene products that are involved in the disease process, much less identifying those which may serve as therapeutic targets for the diagnosis and treatment of various forms of cardiovascular disease.

3. SUMMARY OF THE INVENTION

The present invention relates to methods and compositions for the treatment and diagnosis of cardiovascular disease, including but not limited to, atherosclerosis, ischemia/reperfusion, hypertension, restenosis, and arterial inflammation. Specifically, genes are identified and described which are differentially expressed in cardiovascular disease states, relative to their expression in normal, or non-cardiovascular disease states.

The present invention further relates to screening methods to identify compositions and their therapeutic use for the treatment of fibroproliferative and oncogenic disorders, including diabetic retinopathy, cancer, tumorigenesis, vascularization of tumors, angiogenesis artherosclerosis inflammation and fibrosis.

"Differential expression", as used herein, refers to both quantitative as well as qualitative differences in the genes' temporal and/or tissue expression patterns. Differentially expressed genes may represent "fingerprint genes," and/or "target genes." "Fingerprint gene," as used

herein, refers to a differentially expressed gene whose expression pattern may be utilized as part of a prognostic or diagnostic cardiovascular disease evaluation, or which, alternatively, may be used in methods for identifying compounds useful for the treatment of cardiovascular disease. "Target gene", as used herein, refers to a differentially expressed gene involved in cardiovascular disease such that modulation of the level of target gene expression or of target gene product activity may act to ameliorate a cardiovascular disease condition. Compounds that modulate target gene expression or activity of the target gene product can be used in the treatment of cardiovascular disease.

Further, "pathway genes" are defined via the ability of their products to interact with other gene products involved in cardiovascular disease. Pathway genes may also exhibit target gene and/or fingerprint gene characteristics. Although the genes described herein may be differentially expressed with respect to cardiovascular disease, and/or their products may interact with gene products important to cardiovascular disease, the genes may also be involved in mechanisms important to additional cardiovascular processes.

The invention includes the products of such fingerprint, target, and pathway genes, as well as antibodies to such gene products. Furthermore, the engineering and use of cell- and animal-based models of cardiovascular disease to which such gene products may contribute are also described.

The present invention encompasses methods for prognostic and diagnostic evaluation of cardiovascular disease conditions, and for the identification of subjects exhibiting a predisposition to such conditions. Furthermore, the invention provides methods for evaluating the efficacy of drugs, and monitoring the progress of patients, involved in clinical trials for the treatment of cardiovascular disease.

The invention also provides methods for the identification of compounds that modulate the expression of genes or the activity of gene products involved in

cardiovascular disease, as well as methods for the treatment of cardiovascular disease which may involve the administration of such compounds to individuals exhibiting cardiovascular disease symptoms or tendencies.

The invention also provides methods for the identification of compounds that modulate the expression of genes or the activity of gene products involved in fibroproliferative or oncogenic disorders, including tumorigenesis and the vascularization of tumors.

In addition, the invention encompasses methods for treating cardiovascular disease and fibroproliferative or oncogenic disorders by administering compounds and other substances that modulate the overall activity of the target gene products. Compounds and other substances can effect such modulation either on the level of target gene expression or target protein activity.

The invention is based, in part, on systematic search strategies involving in vivo and in vitro cardiovascular disease paradigms coupled with sensitive and high throughput gene expression assays. In contrast to approaches that merely evaluate the expression of a given gene product presumed to play a role in a disease process, the search strategies and assays used herein permit the identification of all genes, whether known or novel, that are expressed or repressed in the disease condition, as well as the evaluation of their temporal regulation and function during disease progression. This comprehensive approach and evaluation permits the discovery of novel genes and gene products, as well as the identification of an array of genes and gene products (whether novel or known) involved in novel pathways that play a major role in the disease pathology. Thus, the invention allows one to define targets useful for diagnosis, monitoring, rational drug screening and design, and/or other therapeutic intervention.

In the working examples described herein, eight novel human genes are identified that are demonstrated to be differentially expressed in different cardiovascular disease

states. Additionally, the differential expression of four previously identified human genes is described. The identification of these genes and the characterization of their expression in particular disease states provide newly identified roles in cardiovascular disease for both the novel genes and the known genes.

Bcl-2 and glutathione peroxidase are the products of known genes that are shown herein to be down regulated in monocytes of patients exposed to an atherogenic high fat/high cholesterol diet. Furthermore, counteracting the down-regulation of bcl-2 under atherogenic conditions, as described herein, may ameliorate atherosclerosis. Accordingly, methods are provided for the diagnosis, monitoring in clinical trials, and treatment of cardiovascular disease based upon the discoveries herein regarding the expression patterns of bcl-2 and glutathione peroxidase. Because these two genes were known to be involved in preventing apoptosis, the discovery of their down-regulation under atherogenic conditions provides a novel, positive correlation between apoptosis and atherogenesis. Accordingly, methods provided herein for diagnosing, monitoring, and treating cardiovascular disease may also be based on a number of genes involved in the apoptotic pathway, including but not limited to ICE (IL-1 converting enzyme); Bad; BAG-1 (Bcl-2 associated athanogene 1, Takayama et al., 1995, Cell 80: 279-284); BAX (Bcl-2 associated X protein, Oltvai et al., 1993, Cell 74: 609-619); $BclX_L$ (Boise, et al., 1993, Cell 74: 597-608); BAK (Bcl-2 antagonist killer, Farrow et al., 1995. Nature 374: 631-733); and Bcl-X_S (Tsujmoto et al., 1984, Science 226: 1097-1099). The cardiovascular diseases that may be so diagnosed, monitored in clinical trials, and treated include but are not limited to atherosclerosis, ischemia/reperfusion, and restenosis.

rchd005, rchd024, rchd032, and rchd036 are newly identified genes that are each up-regulated in endothelial cells treated with IL-1. Accordingly, methods are provided

for the diagnosis, monitoring in clinical trials, and treatment of cardiovascular disease based upon the discoveries herein regarding the expression patterns of rchd005, rchd024, rchd032, and rchd036.

Cyclooxygenase II (COX II), also known as endoperoxide synthase, and Manganese Superoxide Dismutase (MnSOD) are known genes, and rchd502, rchd523, rchd528, and rchd534 are newly identified genes, that are each up-regulated in endothelial cells subjected to shear stress. Accordingly, methods are provided for the diagnosis, monitoring in clinical trials, screening for therapeutically effective compounds, and treatment of cardiovascular disease based upon the discoveries herein regarding the expression patterns of COX II, MnSOD, rchd502, rchd523, rchd528, and rchd534.

More specifically, each of these genes is up-regulated either by IL-1 (rchd005, rchd024, rchd032, and rchd036) or by shear stress (COX II, MnSOD, rchd502, rchd523, rchd528, and rchd534). For those genes that have a causative effect on the disease conditions treatment methods can be designed to reduce or eliminate their expression, particularly in endothelial cells. Alternatively, treatment methods include inhibiting the activity of the protein products of these genes. For those genes that have a protective effect in responding to disease conditions, treatment methods can be designed for enhancing the activity of the products of such genes.

In either situation, detecting expression of these genes in excess of normal expression provides for the diagnosis of cardiovascular disease. Furthermore, in testing the efficacy of compounds during clinical trials, a decrease in the level of the expression of these genes corresponds to a return from a disease condition to a normal state, and thereby indicates a positive effect of the compound. The cardiovascular diseases that may be so diagnosed, monitored in clinical trials, and treated include but are not limited

to atherosclerosis, ischemia/reperfusion, hypertension, restenosis, and arterial inflammation.

Membrane bound target gene products containing extracellular domains can be a particularly useful target for treatment methods as well as diagnostic and clinical monitoring methods. The rchd523 gene, for example, encodes a transmembrane protein, which contains seven transmembrane domains and, therefore, can be readily contacted by other compounds on the cell surface. Accordingly, natural ligands, derivatives of natural ligands, and antibodies that bind to the rchd523 gene product can be utilized to inhibit its activity, or alternatively, to target the specific destruction of cells that are in the disease state. Furthermore, the extracellular domains of the rchd523 gene product provide especially efficient screening systems for identifying compounds that bind to the rchd523 gene product. Compounds that bind the receptor domain of the rchd523 gene product, for example, can be identified by their ability to mobilize Ca^{2+} and thereby produce a fluorescent signal, as described in Section 5.5.1, below.

Such an assay system can also be used to screen and identify antagonists of the interaction between the rchd523 gene product and ligands that bind to the rchd523 gene product. For example, the compounds can compete with the endogenous (i.e., natural) ligand for the rchd523 gene product. The resulting reduction in the amount of ligand-bound rchd523 gene transmembrane protein will modulate the activity of disease state cells, such as endothelial cells. Soluble proteins or peptides, such as peptides comprising one or more of the extracellular domains, or portions and/or analogs thereof of the rchd523 gene product, including, for example, soluble fusion proteins such as Ig-tailed fusion proteins, can be particularly useful for this purpose.

Similarly, antibodies that are specific to one or more of the extracellular domains of the rchd523 product provide for the ready detection of this target gene product in diagnostic tests or in clinical test monitoring.

Accordingly, endothelial cells can be treated, either in vivo or in vitro, with such a labeled antibody to determine the disease state of endothelial cells. Because the rchd523 gene product is up-regulated in endothelial cells under shear stress, its detection positively corresponds with cardiovascular disease.

Such methods for treatment, diagnosis, and clinical test monitoring which use the rchd523 gene product as described above can also be applied to other target genes that encode transmembrane gene products, including but not limited to rchd502, which each contains 12 transmembrane domains, and rchd528, which contains one transmembrane domain in addition to its extracellular domain.

The invention is based in part on the identification of novel protein-protein interactions of the rchd534 protein with itself and with the fchd540 protein, encoded by the fchd540 gene, as well as interactions of the rchd534 protein or the fchd540 protein with other protein members of the TGF-β signalling pathway. The fchd540 gene and protein were described in co-pending Application No. 08/799,910, filed February 13, 1997, which is hereby incorporated by reference in its entirety. Screening methods are provided for identifying compounds and other substances for treating cardiovascular disease by assaying their ability to inhibit these interactions. Furthermore, methods are provided for identifying compounds and other substances that enhance the TGF-β response by modulating the expression of the rchd534 or fchd540 genes or the activity of their gene products. In addition, methods are provided for treating cardiovascular disease by administering compounds and other substances that inhibit these protein interactions.

In addition, the invention is based in part on the identification of the endothelial cell specific expression pattern of the rchd534 and fchd540 genes, whose protein products inhibit the TGF-β response. The fchd540 gene has been mapped to regions of the human genome that have been implicated in the pathogenesis of several human malignancies.

The invention is further based on the finding that these genes and mutants thereof may be used to modulate TGF-β induced signalling in endothelial cells. Accordingly, the rchd534 and fchd540 genes may be targets for intervention in a variety of inflammatory and fibroproliferative disorders that involve endothelial cells, including, but not limited to, oncology related disorders, disorders related to vascularization, such as cancer angiogenesis, inflammation, and fibrosis.

Both fchd540 and rchd534 are up-regulated in response to laminar shear stress and are specifically expressed in vascular tissue. These findings combined with the observations that both the fchd540 and rchd534 proteins specifically inhibit TGF-β signalling and that the fchd540 and rchd534 genes are located in an area of the human genome implicated in the pathogenesis of several human malignancies indicates that they are excellent and specific targets for therapeutic intervention in the treatment of fibroproliferative and oncogenic disorders including tumorigenesis and vascularization.

The invention is also based in part on the discovery that the rchd534 gene encodes a second protein, in addition to the rchd534 protein described in co-pending application number 08/485,573. Specifically, the rchd534 gene also encodes the rchd534-long protein. The rchd534 protein and the rchd534-long protein are encoded by alternative spliceoforms of rchd534 mRNA. Thus, the rchd534 protein is encoded by the short rchd534 spliceoform, and the rchd534-long protein is encoded by the long rchd534 spliceoform. The rchd534 and rchd534-long protein each has a common C-terminal domain (MH2 domain). The rchd534-long protein contains an N-terminal MH1 domain that is not present in the rchd534 protein. Like rchd534 and fchd540, the rchd534-long protein inhibits TGF-β signalling.

The examples presented in Sections 6-9, below, demonstrate the use of the cardiovascular disease paradigms

of the invention to identify cardiovascular disease target genes.

The example presented in Section 10, below, demonstrates the use of fingerprint genes in diagnostics and as surrogate markers for testing the efficacy of candidate drugs in basic research and in clinical trials.

The example presented in Section 11, below, demonstrates the use of fingerprint genes, particularly rchd523, in the imaging of a diseased cardiovascular tissue.

The example presented in Section 12, below, demonstrates the use of target genes, particularly rchd523, in screening for ligands of target gene product receptor domains, as well as antagonists of the ligand-receptor interaction.

The example presented in Section 15, below, demonstrates the interaction of two target gene products, the rchd534 and fchd540 proteins, and the further characterization of their roles in oncology, angiogenesis, cardiovascular disease and the TGF-β signalling pathway.

The example presented in Section 16, below, describes the discovery of the novel rchd534-long protein and demonstrates that the rchd534-long protein inhibits TGF-β signalling, as do the rchd534 and fchd540 proteins.

4. DESCRIPTION OF THE FIGURES

FIG.1. Band rchd005 DNA sequence. The sequence was determined by sequencing the insert of pRCHD005, resulting from the ligation of amplified rchd005 sequences into the TA cloning vector.

FIG.2. Band rchd024 DNA sequence. The sequence was determined by sequencing the insert of pRCHD024, resulting from the ligation of amplified rchd024 sequences into the TA cloning vector.

FIG.3. Band rchd032 DNA sequence. The sequence was determined by sequencing the insert of pRCHD032, resulting from the ligation of amplified rchd032 sequences into the TA cloning vector.

FIG.4. Band rchd036 DNA sequence. The sequence was determined by sequencing the insert of pRCHD036, resulting from the ligation of amplified rchd036 sequences into the TA cloning vector.

FIG.5. DNA and encoded amino acid sequence of the rchd502 gene.

FIG.6. DNA and encoded amino acid sequence of the rchd523 gene.

FIG.7. DNA and encoded amino acid sequence of the rchd528 gene.

FIG.8. DNA and encoded amino acid sequence of the rchd534 cDNA encoding the rchd534 protein.

FIG. 9. DNA and encoded amino acid sequence of the rchd534-long cDNA encoding the rchd534-long protein.

FIG.10. DNA and encoded amino acid sequence of the fchd540 gene.

FIG.11. Schematic comparison of the rchd534 protein and the rchd534-long protein.

FIG.12. Northern blot analysis of expression of rchd534 mRNA (short spliceoform) under shear stress. RNA was prepared from HUVEC's that were untreated (control) and treated with shear stress for 1 hr. and 6 hr. The blot was probed with labeled rchd534 DNA.

5. DETAILED DESCRIPTION OF THE INVENTION

Methods and compositions for the diagnosis and treatment of cardiovascular disease, including but not limited to atherosclerosis, ischemia/reperfusion, hypertension, restenosis, and arterial inflammation, are described. Methods and compositions for the treatment of oncogenic related disorders, including tumorigenesis and the vascularization of tumors, are also described. The invention is based, in part, on the evaluation of the expression and role of all genes that are differentially expressed in paradigms that are physiologically relevant to the disease condition. This permits the definition of disease pathways

and the identification of targets in the pathway that are useful both diagnostically and therapeutically.

Genes, termed "target genes" and/or "fingerprint genes" which are differentially expressed in cardiovascular disease conditions, relative to their expression in normal, or non-cardiovascular disease conditions, are described in Section 5.4. Additionally, genes, termed "pathway genes" whose gene products exhibit an ability to interact with gene products involved in cardiovascular disease are also described in Section 5.4. Pathway genes may additionally have fingerprint and/or target gene characteristics. Methods for the identification of such fingerprint, target, and pathway genes are described in Sections 5.1, 5.2, and 5.3.

Further, the gene products of such fingerprint, target, and pathway genes are described in Section 5.4.2, antibodies to such gene products are described in Section 5.4.3, as are cell- and animal-based models of cardiovascular disease and fibroproliferative and oncogenic related disorders to which such gene products may contribute, in Section 5.4.4.

Methods for the identification of compounds which modulate the expression of genes or the activity of gene products involved in cardiovascular disease and fibroproliferative and oncogenic related disorders are described in Section 5.5. Methods for monitoring the efficacy of compounds during clinical trials are described in Section 5.5.4. Additionally described below, in Section 5.6, are methods for the treatment of cardiovascular disease.

Also discussed below, in Section 5.8, are methods for prognostic and diagnostic evaluation of cardiovascular disease and fibroproliferative and oncogenic related disorders, including the identification of subjects exhibiting a predisposition to this disease, and the imaging of cardiovascular disease conditions.

5.1. IDENTIFICATION OF DIFFERENTIALLY EXPRESSED GENES

This section describes methods for the identification of genes which are involved in cardiovascular disease, including but not limited to atherosclerosis, ischemia/reperfusion, hypertension, restenosis, and arterial inflammation. Such genes may represent genes which are differentially expressed in cardiovascular disease conditions relative to their expression in normal, or non-cardiovascular disease conditions. Such differentially expressed genes may represent "target" and/or "fingerprint" genes. Methods for the identification of such differentially expressed genes are described, below, in this section. Methods for the further characterization of such differentially expressed genes, and for their identification as target and/or fingerprint genes, are presented, below, in Section 5.3.

"Differential expression" as used herein refers to both quantitative as well as qualitative differences in the genes' temporal and/or tissue expression patterns. Thus, a differentially expressed gene may have its expression activated or completely inactivated in normal versus cardiovascular disease conditions (e.g., treated with oxidized LDL versus untreated), or under control versus experimental conditions. Such a qualitatively regulated gene will exhibit an expression pattern within a given tissue or cell type which is detectable in either control or cardiovascular disease subjects, but is not detectable in both. Alternatively, such a qualitatively regulated gene will exhibit an expression pattern within a given tissue or cell type which is detectable in either control or experimental subjects, but is not detectable in both. "Detectable", as used herein, refers to an RNA expression pattern which is detectable via the standard techniques of differential display, reverse transcriptase- (RT-) PCR and/or Northern analyses, which are well known to those of skill in the art.

Alternatively, a differentially expressed gene may have its expression modulated, i.e., quantitatively increased

or decreased, in normal versus cardiovascular disease states, or under control versus experimental conditions. The degree to which expression differs in normal versus cardiovascular disease or control versus experimental states need only be large enough to be visualized via standard characterization techniques, such as, for example, the differential display technique described below. Other such standard characterization techniques by which expression differences may be visualized include but are not limited to quantitative RT-PCR and Northern analyses.

Differentially expressed genes may be further described as target genes and/or fingerprint genes. "Fingerprint gene," as used herein, refers to a differentially expressed gene whose expression pattern may be utilized as part of a prognostic or diagnostic cardiovascular disease evaluation, or which, alternatively, may be used in methods for identifying compounds useful for the treatment of cardiovascular disease. A fingerprint gene may also have the characteristics of a target gene.

"Target gene", as used herein, refers to a differentially expressed gene involved in cardiovascular disease in a manner by which modulation of the level of target gene expression or of target gene product activity may act to ameliorate symptoms of cardiovascular disease. A target gene may also have the characteristics of a fingerprint gene.

A variety of methods may be utilized for the identification of genes which are involved in cardiovascular disease. These methods include but are not limited to the experimental paradigms described, below, in Section 5.1.1. Material from the paradigms may be characterized for the presence of differentially expressed gene sequences as discussed, below, in Section 5.1.2.

5.1.1. PARADIGMS FOR THE IDENTIFICATION OF DIFFERENTIALLY EXPRESSED GENES

One strategy for identifying genes that are involved in cardiovascular disease is to detect genes that are expressed differentially under conditions associated with the disease versus non-disease conditions. The sub-sections below describe a number of experimental systems, called paradigms, which may be used to detect such differentially expressed genes. In general, the paradigms include at least one experimental condition in which subjects or samples are treated in a manner associated with cardiovascular disease, in addition to at least one experimental control condition lacking such disease associated treatment. Differentially expressed genes are detected, as described herein, below, by comparing the pattern of gene expression between the experimental and control conditions.

Once a particular gene has been identified through the use of one such paradigm, its expression pattern may be further characterized by studying its expression in a different paradigm. A gene may, for example, be regulated one way in a given paradigm (e.g., up-regulation), but may be regulated differently in some other paradigm (e.g., down-regulation). Furthermore, while different genes may have similar expression patterns in one paradigm, their respective expression patterns may differ from one another under a different paradigm. Such use of multiple paradigms may be useful in distinguishing the roles and relative importance of particular genes in cardiovascular disease.

5.1.1.1. FOAM CELL PARADIGM - 1

Among the paradigms which may be utilized for the identification of differentially expressed genes involved in atherosclerosis, for example, are paradigms designed to analyze those genes which may be involved in foam cell formation. Such paradigms may serve to identify genes involved in the differentiation of this cell type, or their uptake of oxidized LDL.

One embodiment of such a paradigm, hereinafter referred to as Paradigm A. First, human blood is drawn and peripheral monocytes are isolated by methods routinely practiced in the art. These human monocytes can then be used immediately or cultured in vitro, using methods routinely practiced in the art, for 5 to 9 days where they develop more macrophage-like characteristics such as the up-regulation of scavenger receptors. These cells are then treated for various lengths of time with agents thought to be involved in foam cell formation. These agents include but are not limited to oxidized LDL, acetylated LDL, lysophosphatidylcholine, and homocysteine. Control monocytes that are untreated or treated with native LDL are grown in parallel. At a certain time after addition of the test agents, the cells are harvested and analyzed for differential expression as described in detail in Section 5.1.2., below. The Example presented in Section 6, below, demonstrates in detail the use of such a foam cell paradigm to identify genes which are differentially expressed in treated versus control cells.

5.1.1.2. FOAM CELL PARADIGM - 2

Alternative paradigms involving monocytes for detecting differentially expressed genes associated with atherosclerosis involve the simulation of the phenomenon of transmigration. When monocytes encounter arterial injury, they adhere to the vascular endothelial layer, transmigrate across this layer, and locate between the endothelium and the layer of smooth muscle cells that ring the artery. This phenomenon can be mimicked in vitro by culturing a layer of endothelial cells isolated, for example, from human umbilical cord. Once the endothelial monolayer forms, monocytes drawn from peripheral blood are cultured on top of the endothelium in the presence and absence of LDL. After several hours, the monocytes transmigrate through the endothelium and develop into foam cells after 3 to 5 days when exposed to LDL. In this system, as in vivo, the endothelial cells carry out the

oxidation of LDL which is then taken up by the monocytes. As described in sub-section 5.1.2. below, the pattern of gene expression can then be compared between these foam cells and untreated monocytes.

5.1.1.3. FOAM CELL PARADIGM - 3

Yet another system includes the third cell type, smooth muscle cell, that plays a critical role in atherogenesis (Navab et al., 1988, J. Clin. Invest., 82: 1853). In this system, a multilayer of human aortic smooth muscle cells was grown on a micropore filter covered with a gel layer of native collagen, and a monolayer of human aortic endothelial cells was grown on top of the collagen layer. Exposure of this coculture to human monocytes in the presence of chemotactic factor rFMLP resulted in monocyte attachment to the endothelial cells followed by migration across the endothelial monolayer into the collagen layer of the subendothelial space. This type of culture can also be treated with LDL to generate foam cells. The foam cells can then be harvested and their pattern of gene expression compared to that of untreated cells as explained below in sub-section 5.1.2.

5.1.1.4. IN VIVO MONOCYTE PARADIGM

An alternative embodiment of such paradigms for the study of monocytes, hereinafter referred to as Paradigm B, involves differential treatment of human subjects through the dietary control of lipid consumption. Such human subjects are held on a low fat/low cholesterol diet for three weeks, at which time blood is drawn, monocytes are isolated according to the methods routinely practiced in the art, and RNA is purified, as described below, in sub-section 5.1.2. These same patients are subsequently switched to a high fat /high cholesterol diet and monocyte RNA is purified again. The patients may also be fed a third, combination diet containing high fat/low cholesterol and monocyte RNA may be purified once again. The order in which patients receive the

diets may be varied. The RNA derived from patients maintained on two of the diets, or on all three diets, may then be compared and analyzed for differential gene expression as, explained below in sub-section 5.1.2.

The Example presented in Section 7, below, demonstrates the use of such an in vivo monocyte paradigm to identify genes which are expressed differentially in monocytes of patients maintained on an atherogenic diet versus their expression under a control diet. Such a paradigm may also be used in conjunction with an in vitro preliminary detection system, as described in Section 7, below.

5.1.1.5. ENDOTHELIAL CELL - IL-1 PARADIGM

In addition to the detection of differential gene expression in monocytes, paradigms focusing on endothelial cells may be used to detect genes involved in cardiovascular disease. In one such paradigm, hereinafter referred to as Paradigm C, human umbilical vein endothelial cells (HUVEC's) are grown in vitro. Experimental cultures are treated with human IL-1β, a factor known to be involved in the inflammatory response, in order to mimic the physiologic conditions involved in the atherosclerotic state. Alternatively experimental HUVEC cultures may be treated with lysophosphatidylcholine, a major phospholipid component of atherogenic lipoproteins or oxidized human LDL. Control cultures are grown in the absence of these compounds.

After a certain period of exposure treatment, experimental and control cells are harvested and analyzed for differential gene expression as described in sub-section 5.1.2, below. The Example presented in Section 8, below, demonstrates the use of such an IL-1 induced endothelial cell paradigm to identify sequences which are differentially expressed in treated versus control cells.

5.1.1.6. ENDOTHELIAL CELL - SHEAR STRESS PARADIGM

In another paradigm involving endothelial cells, hereinafter referred to as Paradigm D, cultures are exposed to fluid shear stress which is thought to be responsible for the prevalence of atherosclerotic lesions in areas of unusual circulatory flow. Unusual blood flow also plays a role in the harmful effects of ischemia/reperfusion, wherein an organ receiving inadequate blood supply is suddenly reperfused with an overabundance of blood when the obstruction is overcome.

Cultured HUVEC monolayers are exposed to laminar sheer stress by rotating the culture in a specialized apparatus containing liquid culture medium (Nagel et al., 1994, J. Clin. Invest. 94: 885-891). Static cultures grown in the same medium serve as controls. After a certain period of exposure to shear stress, experimental and control cells are harvested and analyzed for differential gene expression as described in sub-section 5.1.2, below. The Example presented in Section 9, below, demonstrates the use of such a shear stressed endothelial cell paradigm to identify sequences which are differentially expressed in exposed versus control cells.

In all such paradigms designed to identify genes which are involved in cardiovascular disease, including but not limited to those described above in Sections 5.1.1.1 through 5.1.1.6, compounds such as drugs known to have an ameliorative effect on the disease symptoms may be incorporated into the experimental system. Such compounds may include known therapeutics, as well as compounds that are not useful as therapeutics due to their harmful side effects. Test cells that are cultured as explained in the paradigms described in Sections 5.1.1.1 through 5.1.1.6, for example, may be exposed to one of these compounds and analyzed for differential gene expression with respect to untreated cells, according to the methods described below in Section 5.1.2. In principle, according to the particular paradigm, any cell

type involved in the disease may be treated at any stage of the disease process by these compounds.

Test cells may also be compared to unrelated cells (e.g., fibroblasts) that are also treated with the compound, in order to screen out generic effects on gene expression that might not be related to the disease. Such generic effects might be manifest by changes in gene expression that are common to the test cells and the unrelated cells upon treatment with the compound.

By these methods, the genes and gene products upon which these compounds act can be identified and used in the assays described below to identify novel therapeutic compounds for the treatment of cardiovascular disease.

5.1.2. ANALYSIS OF PARADIGM MATERIAL

In order to identify differentially expressed genes, RNA, either total or mRNA, may be isolated from one or more tissues of the subjects utilized in paradigms such as those described earlier in this Section. RNA samples are obtained from tissues of experimental subjects and from corresponding tissues of control subjects. Any RNA isolation technique which does not select against the isolation of mRNA may be utilized for the purification of such RNA samples. See, for example, Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, N.Y.; and Ausubel, F.M. et al., eds., 1987-1993, Current Protocols in Molecular Biology, John Wiley & Sons, Inc. New York, both of which are incorporated herein by reference in their entirety. Additionally, large numbers of tissue samples may readily be processed using techniques well known to those of skill in the art, such as, for example, the single-step RNA isolation process of Chomczynski, P. (1989, U.S. Patent No. 4,843,155), which is incorporated herein by reference in its entirety.

Transcripts within the collected RNA samples which represent RNA produced by differentially expressed genes may be identified by utilizing a variety of methods which are well known to those of skill in the art. For example,

differential screening (Tedder, T.F. et al., 1988, Proc. Natl. Acad. Sci. USA 85:208-212), subtractive hybridization (Hedrick, S.M. et al., 1984, Nature 308:149-153; Lee, S.W. et al., 1984, Proc. Natl. Acad. Sci. USA 88:2825), and, preferably, differential display (Liang, P., and Pardee, A.B., 1993, U.S. Patent No. 5,262,311, which is incorporated herein by reference in its entirety), may be utilized to identify nucleic acid sequences derived from genes that are differentially expressed.

Differential screening involves the duplicate screening of a cDNA library in which one copy of the library is screened with a total cell cDNA probe corresponding to the mRNA population of one cell type while a duplicate copy of the cDNA library is screened with a total cDNA probe corresponding to the mRNA population of a second cell type. For example, one cDNA probe may correspond to a total cell cDNA probe of a cell type derived from a control subject, while the second cDNA probe may correspond to a total cell cDNA probe of the same cell type derived from an experimental subject. Those clones which hybridize to one probe but not to the other potentially represent clones derived from genes differentially expressed in the cell type of interest in control versus experimental subjects.

Subtractive hybridization techniques generally involve the isolation of mRNA taken from two different sources, e.g., control and experimental tissue, the hybridization of the mRNA or single-stranded cDNA reverse-transcribed from the isolated mRNA, and the removal of all hybridized, and therefore double-stranded, sequences. The remaining non-hybridized, single-stranded cDNAs, potentially represent clones derived from genes that are differentially expressed in the two mRNA sources. Such single-stranded cDNAs are then used as the starting material for the construction of a library comprising clones derived from differentially expressed genes.

The differential display technique describes a procedure, utilizing the well known polymerase chain reaction

(PCR; the experimental embodiment set forth in Mullis, K.B., 1987, U.S. Patent No. 4,683,202) which allows for the identification of sequences derived from genes which are differentially expressed. First, isolated RNA is reverse-transcribed into single-stranded cDNA, utilizing standard techniques which are well known to those of skill in the art. Primers for the reverse transcriptase reaction may include, but are not limited to, oligo dT-containing primers, preferably of the reverse primer type of oligonucleotide described below. Next, this technique uses pairs of PCR primers, as described below, which allow for the amplification of clones representing a random subset of the RNA transcripts present within any given cell. Utilizing different pairs of primers allows each of the mRNA transcripts present in a cell to be amplified. Among such amplified transcripts may be identified those which have been produced from differentially expressed genes.

The reverse oligonucleotide primer of the primer pairs may contain an oligo dT stretch of nucleotides, preferably eleven nucleotides long, at its 5′ end, which hybridizes to the poly(A) tail of mRNA or to the complement of a cDNA reverse transcribed from an mRNA poly(A) tail. Second, in order to increase the specificity of the reverse primer, the primer may contain one or more, preferably two, additional nucleotides at its 3′ end. Because, statistically, only a subset of the mRNA derived sequences present in the sample of interest will hybridize to such primers, the additional nucleotides allow the primers to amplify only a subset of the mRNA derived sequences present in the sample of interest. This is preferred in that it allows more accurate and complete visualization and characterization of each of the bands representing amplified sequences.

The forward primer may contain a nucleotide sequence expected, statistically, to have the ability to hybridize to cDNA sequences derived from the tissues of interest. The nucleotide sequence may be an arbitrary one,

and the length of the forward oligonucleotide primer may range from about 9 to about 13 nucleotides, with about 10 nucleotides being preferred. Arbitrary primer sequences cause the lengths of the amplified partial cDNAs produced to be variable, thus allowing different clones to be separated by using standard denaturing sequencing gel electrophoresis.

PCR reaction conditions should be chosen which optimize amplified product yield and specificity, and, additionally, produce amplified products of lengths which may be resolved utilizing standard gel electrophoresis techniques. Such reaction conditions are well known to those of skill in the art, and important reaction parameters include, for example, length and nucleotide sequence of oligonucleotide primers as discussed above, and annealing and elongation step temperatures and reaction times.

The pattern of clones resulting from the reverse transcription and amplification of the mRNA of two different cell types is displayed via sequencing gel electrophoresis and compared. Differences in the two banding patterns indicate potentially differentially expressed genes.

Once potentially differentially expressed gene sequences have been identified via bulk techniques such as, for example, those described above, the differential expression of such putatively differentially expressed genes should be corroborated. Corroboration may be accomplished via, for example, such well known techniques as Northern analysis and/or RT-PCR.

Upon corroboration, the differentially expressed genes may be further characterized, and may be identified as target and/or fingerprint genes, as discussed, below, in Section 5.3.

Also, amplified sequences of differentially expressed genes obtained through, for example, differential display may be used to isolate full length clones of the corresponding gene. The full length coding portion of the gene may readily be isolated, without undue experimentation, by molecular biological techniques well known in the art.

For example, the isolated differentially expressed amplified fragment may be labeled and used to screen a cDNA library. Alternatively, the labeled fragment may be used to screen a genomic library.

PCR technology may also be utilized to isolate full length cDNA sequences. As described, above, in this Section, the isolated, amplified gene fragments obtained through differential display have 5' terminal ends at some random point within the gene and have 3' terminal ends at a position preferably corresponding to the 3' end of the transcribed portion of the gene. Once nucleotide sequence information from an amplified fragment is obtained, the remainder of the gene (i.e., the 5' end of the gene, when utilizing differential display) may be obtained using, for example, RT-PCR.

In one embodiment of such a procedure for the identification and cloning of full length gene sequences, RNA may be isolated, following standard procedures, from an appropriate tissue or cellular source. A reverse transcription reaction may then be performed on the RNA using an oligonucleotide primer complimentary to the mRNA that corresponds to the amplified fragment, for the priming of first strand synthesis. Because the primer is anti-parallel to the mRNA, extension will proceed toward the 5' end of the mRNA. The resulting RNA/DNA hybrid may then be "tailed" with guanines using a standard terminal transferase reaction, the hybrid may be digested with RNAase H, and second strand synthesis may then be primed with a poly-C primer. Using the two primers, the 5' portion of the gene is amplified using PCR. Sequences obtained may then be isolated and recombined with previously isolated sequences to generate a full-length cDNA of the differentially expressed genes of the invention. For a review of cloning strategies and recombinant DNA techniques, see e.g., Sambrook et al., 1989, *supra*; and Ausubel et al., 1989, *supra*.

5.2. IDENTIFICATION OF PATHWAY GENES

This section describes methods for the identification of genes, termed "pathway genes", involved in cardiovascular disease. "Pathway gene", as used herein, refers to a gene whose gene product exhibits the ability to interact with gene products involved in cardiovascular disease. A pathway gene may be differentially expressed and, therefore, may additionally have the characteristics of a target and/or fingerprint gene.

Any method suitable for detecting protein-protein interactions may be employed for identifying pathway gene products by identifying interactions between gene products and gene products known to be involved in cardiovascular disease. Such known gene products may be cellular or extracellular proteins. Those gene products which interact with such known gene products represent pathway gene products and the genes which encode them represent pathway genes.

Among the traditional methods which may be employed are co-immunoprecipitation, crosslinking and co-purification through gradients or chromatographic columns. Utilizing procedures such as these allows for the identification of pathway gene products. Once identified, a pathway gene product may be used, in conjunction with standard techniques, to identify its corresponding pathway gene. For example, at least a portion of the amino acid sequence of the pathway gene product may be ascertained using techniques well known to those of skill in the art, such as via the Edman degradation technique (see, e.g., Creighton, 1983, Proteins: Structures and Molecular Principles, W.H. Freeman & Co., N.Y., pp.34-49). The amino acid sequence obtained may be used as a guide for the generation of oligonucleotide mixtures that can be used to screen for pathway gene sequences. Screening made be accomplished, for example by standard hybridization or PCR techniques. Techniques for the generation of oligonucleotide mixtures and screening are well-known. (See, e.g., Ausubel, *supra*., and PCR Protocols:

A Guide to Methods and Applications, 1990, Innis, M. et al., eds. Academic Press, Inc., New York).

Additionally, methods may be employed which result in the simultaneous identification of pathway genes which encode the protein interacting with a protein involved in cardiovascular disease. These methods include, for example, probing expression libraries with labeled protein known or suggested to be involved in cardiovascular disease, using this protein in a manner similar to the well known technique of antibody probing of λgt11 libraries.

One such method which detects protein interactions *in vivo*, the two-hybrid system, is described in detail for illustration only and not by way of limitation. One version of this system has been described (Chien et al., 1991, Proc. Natl. Acad. Sci. USA, 88:9578-9582) and is commercially available from Clontech (Palo Alto, CA).

Briefly, utilizing such a system, plasmids are constructed that encode two hybrid proteins: one consists of the DNA-binding domain of a transcription activator protein fused to a known protein, and the other consists of the activator protein's activation domain fused to an unknown protein that is encoded by a cDNA which has been recombined into this plasmid as part of a cDNA library. The plasmids are transformed into a strain of the yeast Saccharomyces cerevisiae that contains a reporter gene (e.g., lacZ) whose regulatory region contains the activator's binding sites. Either hybrid protein alone cannot activate transcription of the reporter gene, the DNA-binding domain hybrid because it does not provide activation function and the activation domain hybrid because it cannot localize to the activator's binding sites. Interaction of the two proteins reconstitutes the functional activator protein and results in expression of the reporter gene, which is detected by an assay for the reporter gene product.

The two-hybrid system or related methodology may be used to screen activation domain libraries for proteins that interact with a known "bait" gene protein. Total genomic or

cDNA sequences may be fused to the DNA encoding an activation domain. Such a library and a plasmid encoding a hybrid of the bait gene protein fused to the DNA-binding domain may be cotransformed into a yeast reporter strain, and the resulting transformants may be screened for those that express the reporter gene. These colonies may be purified and the library plasmids responsible for reporter gene expression may be isolated. DNA sequencing may then be used to identify the proteins encoded by the library plasmids.

For example, and not by way of limitation, the bait gene may be cloned into a vector such that it is translationally fused to the DNA encoding the DNA-binding domain of the GAL4 protein. Also by way of example, for the isolation of genes involved in cardiovascular disease, previously isolated genes known or suggested to play a part in cardiovascular disease may be used as the bait genes. These include but are not limited to the genes for bFGF, IGF-I, VEGF, IL-1, M-CSF, TGFβ, TGFα, TNFα, HB-EGF, PDGF, IFN-γ, and GM-CSF, to name a few.

A cDNA library of the cell line from which proteins that interact with bait gene are to be detected can be made using methods routinely practiced in the art. According to the particular system described herein, for example, the cDNA fragments may be inserted into a vector such that they are translationally fused to the activation domain of GAL4. This library may be co-transformed along with the bait gene-GAL4 fusion plasmid into a yeast strain which contains a *lacZ* gene driven by a promoter which contains the GAL4 activation sequence. A cDNA encoded protein, fused to the GAL4 activation domain, that interacts with bait gene will reconstitute an active GAL4 protein and thereby drive expression of the *lacZ* gene. Colonies which express *lacZ* may be detected by their blue color in the presence of X-gal. The cDNA may then be purified from these strains, and used to produce and isolate the bait gene-interacting protein using techniques routinely practiced in the art.

Once a pathway gene has been identified and isolated, it may be further characterized as, for example, discussed below, in Section 5.3.

5.3. CHARACTERIZATION OF DIFFERENTIALLY EXPRESSED AND PATHWAY GENES

Differentially expressed genes, such as those identified via the methods discussed, above, in Section 5.1.1, pathway genes, such as those identified via the methods discussed, above, in Section 5.2, as well as genes identified by alternative means, may be further characterized by utilizing, for example, methods such as those discussed herein. Such genes will be referred to herein as "identified genes".

Analyses such as those described herein will yield information regarding the biological function of the identified genes. An assessment of the biological function of the differentially expressed genes, in addition, will allow for their designation as target and/or fingerprint genes. Specifically, any of the differentially expressed genes whose further characterization indicates that a modulation of the gene's expression or a modulation of the gene product's activity may ameliorate cardiovascular disease will be designated "target genes", as defined, above, in Section 5.1. Such target genes and target gene products, along with those discussed below, will constitute the focus of the compound discovery strategies discussed, below, in Section 5.5.

Any of the differentially expressed genes whose further characterization indicates that such modulations may not positively affect cardiovascular disease, but whose expression pattern contributes to a gene expression "fingerprint pattern" correlative of, for example, a cardiovascular disease condition will be designated a "fingerprint gene". "Fingerprint patterns" will be more fully discussed, below, in Section 5.8. It should be noted that each of the target genes may also function as

fingerprint genes, as may all or a subset of the pathway genes.

It should further be noted that the pathway genes may also be characterized according to techniques such as those described herein. Those pathway genes which yield information indicating that they are differentially expressed and that modulation of the gene's expression or a modulation of the gene product's activity may ameliorate cardiovascular disease will be also be designated "target genes". Such target genes and target gene products, along with those discussed above, will constitute the focus of the compound discovery strategies discussed, below, in Section 5.5.

It should be additionally noted that the characterization of one or more of the pathway genes may reveal a lack of differential expression, but evidence that modulation of the gene's activity or expression may, nonetheless, ameliorate cardiovascular disease symptoms. In such cases, these genes and gene products would also be considered a focus of the compound discovery strategies of Section 5.5, below.

In instances wherein a pathway gene's characterization indicates that modulation of gene expression or gene product activity may not positively affect cardiovascular disease, but whose expression is differentially expressed and which contributes to a gene expression fingerprint pattern correlative of, for example, a cardiovascular disease state, such pathway genes may additionally be designated as fingerprint genes.

Among the techniques whereby the identified genes may be further characterized, the nucleotide sequence of the identified genes, which may be obtained by utilizing standard techniques well known to those of skill in the art, may be used to further characterize such genes. For example, the sequence of the identified genes may reveal homologies to one or more known sequence motifs which may yield information regarding the biological function of the identified gene product.

Second, an analysis of the tissue distribution of the mRNA produced by the identified genes may be conducted, utilizing standard techniques well known to those of skill in the art. Such techniques may include, for example, Northern analyses and RT-PCR. Such analyses provide information as to whether the identified genes are expressed in tissues expected to contribute to cardiovascular disease. Such analyses may also provide quantitative information regarding steady state mRNA regulation, yielding data concerning which of the identified genes exhibits a high level of regulation in, preferably, tissues which may be expected to contribute to cardiovascular disease.

Such analyses may also be performed on an isolated cell population of a particular cell type derived from a given tissue. Additionally, standard in situ hybridization techniques may be utilized to provide information regarding which cells within a given tissue express the identified gene. Such analyses may provide information regarding the biological function of an identified gene relative to cardiovascular disease in instances wherein only a subset of the cells within the tissue is thought to be relevant to cardiovascular disease.

Such an in situ hybridization analysis is described in the example in Section 14, below. Specifically, the roles of the rchd502 and rchd528 genes in cardiovascular disease were further demonstrated by detecting high levels of their expression specifically within the endothelial cells of diseased tissue removed from a human cardiovascular disease patient, and not in any other cell type present in the tissue, including smooth muscle cells and macrophages. These results clearly demonstrate how detection of differentially expressed genes in the paradigms described herein leads to biologically relevant, novel, specific targets for the treatment and diagnosis of cardiovascular disease.

Third, the sequences of the identified genes may be used, utilizing standard techniques, to place the genes onto genetic maps, e.g., mouse (Copeland & Jenkins, 1991, Trends

in Genetics 7: 113-118) and human genetic maps (Cohen, et al., 1993, Nature 366: 698-701). Such mapping information may yield information regarding the genes' importance to human disease by, for example, identifying genes which map near genetic regions to which known genetic cardiovascular disease tendencies map.

Fourth, the biological function of the identified genes may be more directly assessed by utilizing relevant in vivo and in vitro systems. In vivo systems may include, but are not limited to, animal systems which naturally exhibit cardiovascular disease predisposition, or ones which have been engineered to exhibit such symptoms, including but not limited to the apoE-deficient atherosclerosis mouse model (Plump et al., 1992, Cell 71: 343-353). Such systems are discussed in Section 5.4.4.1, below.

The use of such an in vivo system is described in detail in the example provided in Section 7, below, confirming the role of the target gene bcl-2 (see Table 1, in Section 5.4.1, below). Briefly, bcl-2 expression first was shown to be down-regulated in the apoE-deficient atherosclerosis mouse model. Then, a transgenic mouse was engineered bearing the human bcl-2 gene under the control of a promoter which is induced in monocyte foam cells under atherogenic conditions. To test the effect of the induction of bcl-2 under such conditions, the transgenic mouse is crossed with the apoE-deficient mouse. apoE-deficient progeny bearing the highly expressible bcl-2 gene are then examined for plaque formation and development. Reduction in plaque formation and development in these progeny confirms the effectiveness of intervening in cardiovascular disease through this target gene.

In vitro systems may include, but are not limited to, cell-based systems comprising cell types known or suspected of involvement in cardiovascular disease. Such systems are discussed in detail, below, in Section 5.4.4.2.

In further characterizing the biological function of the identified genes, the expression of these genes may be

modulated within the in vivo and/or in vitro systems, i.e., either over- or underexpressed, and the subsequent effect on the system then assayed. Alternatively, the activity of the product of the identified gene may be modulated by either increasing or decreasing the level of activity in the in vivo and/or in vitro system of interest, and its subsequent effect then assayed.

The information obtained through such characterizations may suggest relevant methods for the treatment of cardiovascular disease involving the gene of interest. For example, treatment may include a modulation of gene expression and/or gene product activity. Characterization procedures such as those described herein may indicate where such modulation should involve an increase or a decrease in the expression or activity of the gene or gene product of interest.

For example, genes which are up-regulated under disease conditions may be involved in causing or exacerbating the disease condition. Treatments directed at down-regulating the activity of such harmfully expressed genes will ameliorate the disease condition. On the other hand, the up-regulation of genes under disease conditions may be part of a protective response by affected cells. Treatments directed at increasing or enhancing the activity of such up-regulated gene products, especially in individuals lacking normal up-regulation, will similarly ameliorate disease conditions. Such methods of treatment are discussed, below, in Section 5.6.

5.4. DIFFERENTIALLY EXPRESSED AND PATHWAY GENES

Identified genes, which include but are not limited to differentially expressed genes such as those identified in Section 5.1.1, above, and pathway genes, such as those identified in Section 5.2, above, are described herein. Specifically, the nucleic acid sequences and gene products of such identified genes are described herein. Further, antibodies directed against the identified genes' products,

and cell- and animal-based models by which the identified genes may be further characterized and utilized are also discussed in this Section.

5.4.1. DIFFERENTIALLY EXPRESSED AND PATHWAY GENE SEQUENCES

The differentially expressed and pathway genes of the invention are listed below, in Table 1. Differentially expressed and pathway gene nucleotide sequences are shown in FIGS. 1-10.

Table 1 lists differentially expressed genes identified through, for example, the paradigms discussed, above, in Section 5.1.1, and below, in the examples presented in Sections 6 through 9 and 16, below. Table 1 also summarizes information regarding the further characterization of such genes.

First, the paradigm used initially to detect the differentially expressed gene is described under the column headed "Paradigm of Original Detection". The expression patterns of those genes which have been shown to be differentially expressed, for example, under one or more of the paradigm conditions described in Section 5.1.1 are summarized under the column headed "Paradigm Expression Pattern". For each of the tested genes, the paradigm which was used and the difference in the expression of the gene among the samples generated is shown. "⇑" indicates that gene expression is up-regulated (i.e., there is an increase in the amount of detectable mRNA) among the samples generated, while "⇓" indicates that gene expression is down-regulated (i.e., there is a decrease in the amount of detectable mRNA) among the samples generated. "Detectable" as used herein, refers to levels of mRNA which are detectable via, for example, standard Northern and/or RT-PCR techniques which are well known to those of skill in the art.

Cell types in which differential expression was detected are also summarized in Table 1 under the column headed "Cell Type Detected in". The column headed

"Chromosomal Location" provides the human chromosome number on which the gene is located. Additionally, in instances wherein the genes contain nucleotide sequences similar or homologous to sequences found in nucleic acid databases, references to such similarities are listed.

The genes listed in Table 1 may be obtained using cloning methods well known to those skilled in the art, including but not limited to the use of appropriate probes to detect the genes within an appropriate cDNA or gDNA (genomic DNA) library. (See, for example, Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratories, which is incorporated by reference herein in its entirety). Probes for the novel sequences reported herein may be obtained directly from the isolated clones deposited with the NRRL or ATCC, as indicated in Table 2, below. Alternatively, oligonucleotide probes for the novel genes may be synthesized based on the DNA sequences disclosed herein in FIGS. 1-10. Such synthetic oligonucleotides may be similarly produced based on the sequences provided for the previously known genes described in the following references: Cleary et al., 1986, Cell 47: 19-28 (bcl-2); Takahashi et al., 1990, J. Biochem 108: 145-148 (glutathione peroxidase); and Jones et al., 1993, J. Biol. Chem. 268: 9049-9054 (prostaglandin endoperoxide synthase II), each of which is incorporated herein in its entirety.

The sequence obtained from clones containing partial coding sequences or non-coding sequences can be used to obtain the entire coding region by using the RACE method (Chenchik, et al., 1995, CLONTECHniques (X) 1: 5-8; Barnes, 1994, Proc. Natl. Acad. Sci. USA 91: 2216-2220; and Cheng et al., Proc. Natl. Acad. Sci. USA 91: 5695-5699). Oligonucleotides can be designed based on the sequence obtained from the partial clone that can amplify a reverse transcribed mRNA encoding the entire coding sequence. This method was used, as described in the example in Section 9, below, to obtain the entire coding region of the rchd523 gene.

Alternatively, probes can be used to screen cDNA libraries prepared from an appropriate cell or cell line in which the gene is transcribed. For example, the genes described herein that were detected in monocytes may be cloned from a cDNA library prepared from monocytes isolated as described in Section 7.1.1, below. In fact, as described in detail in the example in Section 9, below, this method was applied in order to obtain the entire coding region of the rchd534 gene. Briefly, the up-regulation of this gene was detected, under Paradigm D, in HUVEC's subjected to shear stress. Then, amplified partial sequence of the rchd534 gene was subcloned. The insert was then isolated and used to probe a cDNA library prepared from shear stress treated HUVEC's. A cDNA clone containing the entire rchd534 coding region was detected, isolated, and sequenced.

The genes described herein that were detected in endothelial cells may also be cloned from a cDNA library constructed from endothelial cells isolated as described in Progress in Hemostasis and Thrombosis, Vol. 3, P. Spaet, editor, Grune & Stratton Inc., New York, 1-28. Alternatively, the genes may be retrieved from a human placenta cDNA library (Clontech Laboratories, Palo Alto, CA), according to Takahashi et al., 1990, supra; a HUVEC cDNA library as described in Jones et al. 1993, supra; or an acute lymphoblastic leukemia (SUP-B2) cDNA library as described in Cleary et al., 1986, supra, for example. Genomic DNA libraries can be prepared from any source.

TABLE 1

Differentially Expressed and Pathway Genes

Gene	Seq. ID #	Paradigm of Original Detection	Paradigm Expr. Pattern	Cell Type Detected in	Chromosomal Location	Ref	FIG.
Band 14: bcl-2		B	⇓	Monocytes		1	
Glutathione peroxidase		B	⇓	Monocytes		2	
rchd005		C	⇑	Endothelial		New 8	FIG.1
rchd024		C	⇑	Endothelial	4	New	FIG.2
rchd032		C	⇑	Endothelial		New	FIG.3
rchd036		C	⇑	Endothelial	15	New	FIG.4
rchd502		D	⇑	Endothelial		New 4	FIG.5
rchd505: COX II		D	⇑	Endothelial		5	
rchd523		D	⇑	Endothelial	7	New	FIG.6
rchd528		D	⇑	Endothelial		New 8	FIG.7
rchd530: MnSOD		D	⇑	Endothelial		7	
rchd534		D	⇑	Endothelial	15	New 8	FIG.8
rchd534	rchd534-long					New 8	FIG.9
fchd540		D	⇑	Endothelial	18	New 8	FIG.10

1 Cleary et al., 1986, Cell 47: 19-28.
2 Takahashi et al., 1990, J. Biochem. 108: 145-148.

3 Shark Na-K-Cl cotransporter, Xu et al., 1994 Proc. Natl. Acad. Sci. U.S.A. 91: 2201-2205.
4 Rat matrin F/G, Hakes et al., 1991 Proc. Natl. Acad. Sci. U.S.A. 88: 6186-6190.
5 Jones et al., 1993, J. Biol. Chem. 268: 9049-9054.
6 *Xenopus* Xotch (homolog of *Drosophila* Notch), Coffman et al., 1990, Science 249: 1438-1441.
7 Heckl, 1988, Nucl. Acids Res. 16: 6224.
8 *Drosohila Mothers against dpp (Mad)*, Sekelsky et al., 1995, Genetics 139: 1347-1358.

Table 2, below, lists isolated clones that contain sequences of the novel genes listed in Table 1. Such clones were produced from amplified sequences of the indicated differential display band which were subcloned into the TA cloning vector (Invitrogen, San Diego, CA), as described in Section 6.1, below. Also listed in Table 2, below, are the strains deposited with the NRRL or ATCC which contain each such clone. Such strains were produced by transforming E. coli strain INVαF' (Invitrogen) with the indicated plasmid, as described in Section 6.1, below. The names of the plasmids containing the entire coding region of a novel gene bear the prefix pFCHD, and the names of the strains carrying these plasmids bear the prefix FCHD.

TABLE 2

GENE		Strain Deposited	Plasmid Clone Contained in Deposited Strain
rchd005		RCHD005	pRCHD005
rchd024		RCHD024	pRCHD024
rchd032		RCHD032	pRCHD032
rchd036		RCHD036	pRCHD036
rchd502		FCHD502SF	pFCHD502SF
		FCHD502SJ	pFCHD502SJ
		RCHD502	pRCHD502
rchd523		FCHD523	pFCHD523
		RCDH523	pRCHD523
rchd528		FCHD528A	pFCHD528A
		FCHD528B	pFCHD528B
		FCHD528C	pFCHD528C
		FCHD523	pRCHD528
rchd534	short spliceoform	FCHD534	pFCHD534
	long spliceoform		pHL6TA1A

As used herein, "differentially expressed gene" (*i.e.* target and fingerprint gene) or "pathway gene" refers to (a) a gene containing at least one of the DNA sequences disclosed herein (as shown in FIGS. 1-10), or contained in the clones listed in Table 2, as deposited with the ATCC or NRRL; (b) any DNA sequence that encodes the amino acid sequence encoded by the DNA sequences disclosed herein (as shown in FIGS. 1-10), contained in the clones, listed in Table 2, as deposited with the ATCC or NRRL or contained within the coding region of the gene to which the DNA sequences disclosed herein (as shown in FIGS. 1-10) or contained in the clones listed in Table 2, as deposited with the ATCC or NRRL, belong; (c) any DNA sequence that hybridizes to the complement of the coding sequences disclosed herein, contained in the clones listed in Table 2, as deposited with the ATCC or NRRL, or contained within the coding region of the gene to which the DNA sequences disclosed herein (as shown in FIGS. 1-10) or contained in the clones listed in Table 2, as deposited with the ATCC or NRRL, belong, under highly stringent conditions, e.g., hybridization to filter-bound DNA in 0.5 M $NaHPO_4$, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1xSSC/0.1% SDS at 68°C (Ausubel F.M. et al., eds., 1989, Current Protocols in Molecular Biology, Vol. I, Green Publishing Associates, Inc., and John Wiley & sons, Inc., New York, at p. 2.10.3) and encodes a gene product functionally equivalent to a gene product encoded by sequences contained within the clones listed in Table 2; and/or (d) any DNA sequence that hybridizes to the complement of the coding sequences disclosed herein, (as shown in FIGS. 1-10) contained in the clones listed in Table 2, as deposited with the ATCC or NRRL or contained within the coding region of the gene to which DNA sequences disclosed herein (as shown in FIGS. 1-10) or contained in the clones, listed in Table 2, as deposited with the ATCC or NRRL, belong, under less stringent conditions, such as moderately stringent conditions, e.g., washing in 0.2xSSC/0.1% SDS at 42°C (Ausubel et al., 1989,

supra), yet which still encodes a functionally equivalent gene product.

The invention also includes nucleic acid molecules, preferably DNA molecules, that hybridize to, and are therefore the complements of, the DNA sequences (a) through (c), in the preceding paragraph. Such hybridization conditions may be highly stringent or less highly stringent, as described above. In instances wherein the nucleic acid molecules are deoxyoligonucleotides ("oligos"), highly stringent conditions may refer, e.g., to washing in 6xSSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligos), 48°C (for 17-base oligos), 55°C (for 20-base oligos), and 60°C (for 23-base oligos). These nucleic acid molecules may act as target gene antisense molecules, useful, for example, in target gene regulation and/or as antisense primers in amplification reactions of target gene nucleic acid sequences. Further, such sequences may be used as part of ribozyme and/or triple helix sequences, also useful for target gene regulation. Still further, such molecules may be used as components of diagnostic methods whereby the presence of a cardiovascular disease-causing allele, may be detected.

The invention also encompasses (a) DNA vectors that contain any of the foregoing coding sequences and/or their complements (i.e., antisense); (b) DNA expression vectors that contain any of the foregoing coding sequences operatively associated with a regulatory element that directs the expression of the coding sequences; and (c) genetically engineered host cells that contain any of the foregoing coding sequences operatively associated with a regulatory element that directs the expression of the coding sequences in the host cell. As used herein, regulatory elements include but are not limited to inducible and non-inducible promoters, enhancers, operators and other elements known to those skilled in the art that drive and regulate expression. The invention includes fragments of any of the DNA sequences disclosed herein.

In addition to the gene sequences described above,

homologues of such sequences as may, for example, be present in humans or in other species, may be identified and may be readily isolated, without undue experimentation, by molecular biological techniques well known in the art. Further, there may exist genes at other genetic loci within the genome that encode proteins which have extensive homology to one or more domains of such gene products. These genes may also be identified via similar techniques.

For example, the isolated differentially expressed gene sequence may be labeled and used to screen a cDNA library constructed from mRNA obtained from the organism of interest. Hybridization conditions will be of a lower stringency when the cDNA library was derived from an organism different from the type of organism from which the labeled sequence was derived. Alternatively, the labeled fragment may be used to screen a genomic library derived from the organism of interest, again, using appropriately stringent conditions. Such low stringency conditions will be well known to those of skill in the art, and will vary predictably depending on the specific organisms from which the library and the labeled sequences are derived. For guidance regarding such conditions see, for example, Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, Cold Springs Harbor Press, N.Y.; and Ausubel et al., 1989, Current Protocols in Molecular Biology, Green Publishing Associates and Wiley Interscience, N.Y.

Further, a previously unknown differentially expressed or pathway gene-type sequence may be isolated by performing PCR using two degenerate oligonucleotide primer pools designed on the basis of amino acid sequences within the gene of interest. The template for the reaction may be cDNA obtained by reverse transcription of mRNA prepared from human or non-human cell lines or tissue known or suspected to express a differentially expressed or pathway gene allele.

The PCR product may be subcloned and sequenced to insure that the amplified sequences represent the sequences of a differentially expressed or pathway gene-like nucleic

acid sequence. The PCR fragment may then be used to isolate a full length cDNA clone by a variety of methods. For example, the amplified fragment may be labeled and used to screen a bacteriophage cDNA library. Alternatively, the labeled fragment may be used to screen a genomic library.

PCR technology may also be utilized to isolate full length cDNA sequences. For example, RNA may be isolated, following standard procedures, from an appropriate cellular or tissue source. A reverse transcription reaction may be performed on the RNA using an oligonucleotide primer specific for the most 5' end of the amplified fragment for the priming of first strand synthesis. The resulting RNA/DNA hybrid may then be "tailed" with guanines using a standard terminal transferase reaction, the hybrid may be digested with RNAase H, and second strand synthesis may then be primed with a poly-C primer. Thus, cDNA sequences upstream of the amplified fragment may easily be isolated. For a review of cloning strategies which may be used, see e.g., Sambrook et al., 1989, *supra*.

In cases where the differentially expressed or pathway gene identified is the normal, or wild type, gene, this gene may be used to isolate mutant alleles of the gene. Such an isolation is preferable in processes and disorders which are known or suspected to have a genetic basis. Mutant alleles may be isolated from individuals either known or suspected to have a genotype which contributes to cardiovascular disease symptoms. Mutant alleles and mutant allele products may then be utilized in the therapeutic and diagnostic assay systems described below.

A cDNA of the mutant gene may be isolated, for example, by using PCR, a technique which is well known to those of skill in the art. In this case, the first cDNA strand may be synthesized by hybridizing an oligo-dT oligonucleotide to mRNA isolated from tissue known or suspected to be expressed in an individual putatively carrying the mutant allele, and by extending the new strand with reverse transcriptase. The second strand of the cDNA is

then synthesized using an oligonucleotide that hybridizes specifically to the 5′ end of the normal gene. Using these two primers, the product is then amplified via PCR, cloned into a suitable vector, and subjected to DNA sequence analysis through methods well known to those of skill in the art. By comparing the DNA sequence of the mutant gene to that of the normal gene, the mutation(s) responsible for the loss or alteration of function of the mutant gene product can be ascertained.

Alternatively, a genomic or cDNA library can be constructed and screened using DNA or RNA, respectively, from a tissue known to or suspected of expressing the gene of interest in an individual suspected of or known to carry the mutant allele. The normal gene or any suitable fragment thereof may then be labeled and used as a probed to identify the corresponding mutant allele in the library. The clone containing this gene may then be purified through methods routinely practiced in the art, and subjected to sequence analysis as described, above, in this Section.

Additionally, an expression library can be constructed utilizing DNA isolated from or cDNA synthesized from a tissue known to or suspected of expressing the gene of interest in an individual suspected of or known to carry the mutant allele. In this manner, gene products made by the putatively mutant tissue may be expressed and screened using standard antibody screening techniques in conjunction with antibodies raised against the normal gene product, as described, below, in Section 5.4.3. (For screening techniques, see, for example, Harlow, E. and Lane, eds., 1988, "Antibodies: A Laboratory Manual", Cold Spring Harbor Press, Cold Spring Harbor.) In cases where the mutation results in an expressed gene product with altered function (e.g., as a result of a missense mutation), a polyclonal set of antibodies are likely to cross-react with the mutant gene product. Library clones detected via their reaction with such labeled antibodies can be purified and subjected to sequence analysis as described in this Section, above.

5.4.2. DIFFERENTIALLY EXPRESSED AND PATHWAY GENE PRODUCTS

Differentially expressed and pathway gene products include those proteins encoded by the differentially expressed and pathway gene sequences described in Section 5.4.1, above. Specifically, differentially expressed and pathway gene products may include differentially expressed and pathway gene polypeptides encoded by the differentially expressed and pathway gene sequences contained in the clones listed in Table 2, above, as deposited with the NRRL or ATCC, or contained in the coding regions of the genes to which DNA sequences disclosed herein (in FIGS. 1-10) or contained in the clones, listed in Table 2, as deposited with the NRRL or ATCC, belong, for example.

In addition, differentially expressed and pathway gene products may include proteins that represent functionally equivalent gene products. Such an equivalent differentially expressed or pathway gene product may contain deletions, additions or substitutions of amino acid residues within the amino acid sequence encoded by the differentially expressed or pathway gene sequences described, above, in Section 5.4.1, but which result in a silent change, thus producing a functionally equivalent differentially expressed on pathway gene product. Amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved.

For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Functionally equivalent", as utilized herein, refers to a protein capable of exhibiting a substantially similar in vivo activity as the endogenous differentially expressed or pathway gene products encoded by

the differentially expressed or pathway gene sequences described in Section 5.4.1, above. Alternatively, when utilized as part of assays such as those described, below, in Section 5.5, "functionally equivalent" may refer to peptides capable of interacting with other cellular or extracellular molecules in a manner substantially similar to the way in which the corresponding portion of the endogenous differentially expressed or pathway gene product would.

The differentially expressed or pathway gene products may be produced by recombinant DNA technology using techniques well known in the art. Thus, methods for preparing the differentially expressed or pathway gene polypeptides and peptides of the invention by expressing nucleic acid encoding differentially expressed or pathway gene sequences are described herein. Methods which are well known to those skilled in the art can be used to construct expression vectors containing differentially expressed or pathway gene protein coding sequences and appropriate transcriptional/translational control signals. These methods include, for example, in vitro recombinant DNA techniques, synthetic techniques and in vivo recombination/genetic recombination. See, for example, the techniques described in Sambrook et al., 1989, *supra*, and Ausubel et al., 1989, *supra*. Alternatively, RNA capable of encoding differentially expressed or pathway gene protein sequences may be chemically synthesized using, for example, synthesizers. See, for example, the techniques described in "Oligonucleotide Synthesis", 1984, Gait, M.J. ed., IRL Press, Oxford, which is incorporated by reference herein in its entirety.

A variety of host-expression vector systems may be utilized to express the differentially expressed or pathway gene coding sequences of the invention. Such host-expression systems represent vehicles by which the coding sequences of interest may be produced and subsequently purified, but also represent cells which may, when transformed or transfected with the appropriate nucleotide coding sequences, exhibit the differentially expressed or pathway gene protein of the

invention in situ. These include but are not limited to microorganisms such as bacteria (e.g., *E. coli*, *B. subtilis*) transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing differentially expressed or pathway gene protein coding sequences; yeast (e.g. *Saccharomyces*, *Pichia*) transformed with recombinant yeast expression vectors containing the differentially expressed or pathway gene protein coding sequences; insect cell systems infected with recombinant virus expression vectors (e.g., baculovirus) containing the differentially expressed or pathway gene protein coding sequences; plant cell systems infected with recombinant virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (e.g., Ti plasmid) containing differentially expressed or pathway gene protein coding sequences; or mammalian cell systems (e.g. COS, CHO, BHK, 293, 3T3) harboring recombinant expression constructs containing promoters derived from the genome of mammalian cells (e.g., metallothionein promoter) or from mammalian viruses (e.g., the adenovirus late promoter; the vaccinia virus 7.5K promoter).

In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the differentially expressed or pathway gene protein being expressed. For example, when a large quantity of such a protein is to be produced, for the generation of antibodies or to screen peptide libraries, for example, vectors which direct the expression of high levels of fusion protein products that are readily purified may be desirable. Such vectors include, but are not limited, to the *E. coli* expression vector pUR278 (Ruther et al., 1983, EMBO J. 2:1791), in which the differentially expressed or pathway gene protein coding sequence may be ligated individually into the vector in frame with the *lac Z* coding region so that a fusion protein is produced; pIN vectors (Inouye & Inouye, 1985, Nucleic Acids Res. 13:3101-3109; Van Heeke & Schuster,

1989, J. Biol. Chem. 264:5503-5509); and the like. pGEX vectors may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. The pGEX vectors are designed to include thrombin or factor Xa protease cleavage sites so that the cloned target gene protein can be released from the GST moiety.

In a preferred embodiment, full length cDNA sequences are appended with in-frame Bam HI sites at the amino terminus and Eco RI sites at the carboxyl terminus using standard PCR methodologies (Innis et al., 1990, *supra*) and ligated into the pGEX-2TK vector (Pharmacia, Uppsala, Sweden). The resulting cDNA construct contains a kinase recognition site at the amino terminus for radioactive labelling and glutathione S-transferase sequences at the carboxyl terminus for affinity purification (Nilsson, et al., 1985, EMBO J. 4: 1075; Zabeau and Stanley, 1982, *EMBO J.* 1: 1217.

In an insect system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes. The virus grows in *Spodoptera frugiperda* cells. The differentially expressed or pathway gene coding sequence may be cloned individually into non-essential regions (for example the polyhedrin gene) of the virus and placed under control of an AcNPV promoter (for example the polyhedrin promoter). Successful insertion of differentially expressed or pathway gene coding sequence will result in inactivation of the polyhedrin gene and production of non-occluded recombinant virus (i.e., virus lacking the proteinaceous coat coded for by the polyhedrin gene). These recombinant viruses are then used to infect Spodoptera frugiperda cells in which the inserted gene is expressed. (E.g., see Smith et al., 1983, J. Virol. 46: 584; Smith, U.S. Patent No. 4,215,051).

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, the differentially expressed or pathway gene coding sequence of interest may be ligated to an adenovirus transcription/translation control complex, e.g., the late promoter and tripartite leader sequence. This chimeric gene may then be inserted in the adenovirus genome by in vitro or in vivo recombination. Insertion in a non-essential region of the viral genome (e.g., region E1 or E3) will result in a recombinant virus that is viable and capable of expressing differentially expressed or pathway gene protein in infected hosts. (E.g., See Logan & Shenk, 1984, Proc. Natl. Acad. Sci. USA 81:3655-3659). Specific initiation signals may also be required for efficient translation of inserted differentially expressed or pathway gene coding sequences. These signals include the ATG initiation codon and adjacent sequences. In cases where an entire differentially expressed or pathway gene, including its own initiation codon and adjacent sequences, is inserted into the appropriate expression vector, no additional translational control signals may be needed. However, in cases where only a portion of the differentially expressed or pathway gene coding sequence is inserted, exogenous translational control signals, including, perhaps, the ATG initiation codon, must be provided. Furthermore, the initiation codon must be in phase with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements, transcription terminators, etc. (see Bittner et al., 1987, Methods in Enzymol. 153:516-544).

In a preferred embodiment, cDNA sequences encoding the full-length open reading frames are ligated into pCMVβ replacing the β-galactosidase gene such that cDNA expression

is driven by the CMV promoter (Alam, 1990, Anal. Biochem. 188: 245-254; MacGregor & Caskey, 1989, Nucl. Acids Res. 17: 2365; Norton & Corrin, 1985, Mol. Cell. Biol. 5: 281).

In addition, a host cell strain may be chosen which modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may be important for the function of the protein. Different host cells have characteristic and specific mechanisms for the post-translational processing and modification of proteins. Appropriate cell lines or host systems can be chosen to ensure the correct modification and processing of the foreign protein expressed. To this end, eukaryotic host cells which possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product may be used. Such mammalian host cells include but are not limited to CHO, VERO, BHK, HeLa, COS, MDCK, 293, 3T3, WI38, etc.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines which stably express the differentially expressed or pathway gene protein may be engineered. Rather than using expression vectors which contain viral origins of replication, host cells can be transformed with DNA controlled by appropriate expression control elements (e.g., promoter, enhancer, sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. Following the introduction of the foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines. This method may advantageously be used to engineer cell lines which express the differentially expressed or

pathway gene protein. Such engineered cell lines may be particularly useful in screening and evaluation of compounds that affect the endogenous activity of the differentially expressed or pathway gene protein.

A number of selection systems may be used, including but not limited to the herpes simplex virus thymidine kinase (Wigler, et al., 1977, Cell 11:223), hypoxanthine-guanine phosphoribosyltransferase (Szybalska & Szybalski, 1962, Proc. Natl. Acad. Sci. USA 48:2026), and adenine phosphoribosyltransferase (Lowy, et al., 1980, Cell 22:817) genes can be employed in tk^-, $hgprt^-$ or $aprt^-$ cells, respectively. Also, antimetabolite resistance can be used as the basis of selection for dhfr, which confers resistance to methotrexate (Wigler, et al., 1980, Natl. Acad. Sci. USA 77:3567; O'Hare, et al., 1981, Proc. Natl. Acad. Sci. USA 78:1527); gpt, which confers resistance to mycophenolic acid (Mulligan & Berg, 1981, Proc. Natl. Acad. Sci. USA 78:2072); neo, which confers resistance to the aminoglycoside G-418 (Colberre-Garapin, et al., 1981, J. Mol. Biol. 150:1); and hygro, which confers resistance to hygromycin (Santerre, et al., 1984, Gene 30:147) genes.

An alternative fusion protein system allows for the ready purification of non-denatured fusion proteins expressed in human cell lines (Janknecht, et al., 1991, Proc. Natl. Acad. Sci. USA 88: 8972-8976). In this system, the gene of interest is subcloned into a vaccinia recombination plasmid such that the gene's open reading frame is translationally fused to an amino-terminal tag consisting of six histidine residues. Extracts from cells infected with recombinant vaccinia virus are loaded onto Ni^{2+}·nitriloacetic acid-agarose columns and histidine-tagged proteins are selectively eluted with imidazole-containing buffers.

When used as a component in assay systems such as those described, below, in Section 5.5, the differentially expressed or pathway gene protein may be labeled, either directly or indirectly, to facilitate detection of a complex formed between the differentially expressed or pathway gene

protein and a test substance. Any of a variety of suitable labeling systems may be used including but not limited to radioisotopes such as ^{125}I; enzyme labelling systems that generate a detectable colorimetric signal or light when exposed to substrate; and fluorescent labels.

Where recombinant DNA technology is used to produce the differentially expressed or pathway gene protein for such assay systems, it may be advantageous to engineer fusion proteins that can facilitate labeling, immobilization and/or detection.

Indirect labeling involves the use of a protein, such as a labeled antibody, which specifically binds to either a differentially expressed or pathway gene product. Such antibodies include but are not limited to polyclonal, monoclonal, chimeric, single chain, Fab fragments and fragments produced by an Fab expression library.

5.4.3. DIFFERENTIALLY EXPRESSED OR PATHWAY GENE PRODUCT ANTIBODIES

Described herein are methods for the production of antibodies capable of specifically recognizing one or more differentially expressed or pathway gene epitopes. Such antibodies may include, but are not limited to polyclonal antibodies, monoclonal antibodies (mAbs), humanized or chimeric antibodies, single chain antibodies, Fab fragments, $F(ab')_2$ fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above. Such antibodies may be used, for example, in the detection of a fingerprint, target, or pathway gene in a biological sample, or, alternatively, as a method for the inhibition of abnormal target gene activity. Thus, such antibodies may be utilized as part of cardiovascular disease treatment methods, and/or may be used as part of diagnostic techniques whereby patients may be tested for abnormal levels of fingerprint, target, or pathway gene proteins, or for the presence of abnormal forms of the such proteins.

For the production of antibodies to a differentially expressed or pathway gene, various host animals may be immunized by injection with a differentially expressed or pathway gene protein, or a portion thereof. Such host animals may include but are not limited to rabbits, mice, and rats, to name but a few. Various adjuvants may be used to increase the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and *Corynebacterium parvum*.

In a preferred embodiment, peptide sequences corresponding to amino sequences of target gene products were selected and submitted to Research Genetics (Huntsville, AL) for synthesis and antibody production. Peptides were modified as described (Tam, J.P., 1988, Proc. Natl. Acad. Sci. USA 85: 5409-5413; Tam, J.P., and Zavala, F., 1989, J. Immunol. Methods 124: 53-61; Tam, J.P., and Lu, Y.A., 1989, Proc. Natl. Acad. Sci. USA 86: 9084-9088), emulsified in an equal volume of Freund's adjuvant and injected into rabbits at 3 to 4 subcutaneous dorsal sites for a total volume of 1.0 ml (0.5 mg peptide) per immunization. The animals were boosted after 2 and 6 weeks and bled at weeks 4, 8, and 10. The blood was allowed to clot and serum was collected by centrifugation.

Polyclonal antibodies are heterogeneous populations of antibody molecules derived from the sera of animals immunized with an antigen, such as target gene product, or an antigenic functional derivative thereof. For the production of polyclonal antibodies, host animals such as those described above, may be immunized by injection with differentially expressed or pathway gene product supplemented with adjuvants as also described above.

Monoclonal antibodies, which are homogeneous populations of antibodies to a particular antigen, may be obtained by any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to the hybridoma technique of Kohler and Milstein, (1975, Nature 256:495-497; and U.S. Patent No. 4,376,110), the human B-cell hybridoma technique (Kosbor et al., 1983, Immunology Today 4:72; Cole et al., 1983, Proc. Natl. Acad. Sci. USA 80:2026-2030), and the EBV-hybridoma technique (Cole et al., 1985, Monoclonal Antibodies And Cancer Therapy, Alan R. Liss, Inc., pp. 77-96). Such antibodies may be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD and any subclass thereof. The hybridoma producing the mAb of this invention may be cultivated in vitro or in vivo. Production of high titers of mAbs in vivo makes this the presently preferred method of production.

In addition, techniques developed for the production of "chimeric antibodies" (Morrison et al., 1984, Proc. Natl. Acad. Sci., 81:6851-6855; Neuberger et al., 1984, Nature, 312:604-608; Takeda et al., 1985, Nature, 314:452-454) by splicing the genes from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. A chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a human immunoglobulin constant region.

Alternatively, techniques described for the production of single chain antibodies (U.S. Patent 4,946,778; Bird, 1988, Science 242:423-426; Huston et al., 1988, Proc. Natl. Acad. Sci. USA 85:5879-5883; and Ward et al., 1989, Nature 334:544-546) can be adapted to produce differentially expressed or pathway gene-single chain antibodies. Single chain antibodies are formed by linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain polypeptide.

Antibody fragments which recognize specific epitopes may be generated by known techniques. For example, such fragments include but are not limited to: the $F(ab')_2$ fragments which can be produced by pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of the $F(ab')_2$ fragments. Alternatively, Fab expression libraries may be constructed (Huse et al., 1989, Science, 246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity.

5.4.4. CELL- AND ANIMAL-BASED MODEL SYSTEMS

Described herein are cell- and animal-based systems which act as models for cardiovascular disease. These systems may be used in a variety of applications. For example, the cell- and animal-based model systems may be used to further characterize differentially expressed and pathway genes, as described, above, in Section 5.3. Such further characterization may, for example, indicate that a differentially expressed gene is a target gene. Second, such assays may be utilized as part of screening strategies designed to identify compounds which are capable of ameliorating cardiovascular disease symptoms, as described, below, in Section 5.5.4. Thus, the animal- and cell-based models may be used to identify drugs, pharmaceuticals, therapies and interventions which may be effective in treating cardiovascular disease. In addition, as described in detail, below, in Section 5.7.1, such animal models may be used to determine the LD_{50} and the ED_{50} in animal subjects, and such data can be used to determine the in vivo efficacy of potential cardiovascular disease treatments.

5.4.4.1. ANIMAL-BASED SYSTEMS

Animal-based model systems of cardiovascular disease may include, but are not limited to, non-recombinant and engineered transgenic animals.

Non-recombinant animal models for cardiovascular disease may include, for example, genetic models. Such genetic cardiovascular disease models may include, for example, apoB or apoR deficient pigs (Rapacz, et al., 1986, Science 234:1573-1577) and Watanabe heritable hyperlipidemic (WHHL) rabbits (Kita et al., 1987, Proc. Natl. Acad. Sci USA 84: 5928-5931).

Non-recombinant, non-genetic animal models of atherosclerosis may include, for example, pig, rabbit, or rat models in which the animal has been exposed to either chemical wounding through dietary supplementation of LDL, or mechanical wounding through balloon catheter angioplasty, for example.

Additionally, animal models exhibiting cardiovascular disease symptoms may be engineered by utilizing, for example, target gene sequences such as those described, above, in Section 5.4.1, in conjunction with techniques for producing transgenic animals that are well known to those of skill in the art. For example, target gene sequences may be introduced into, and overexpressed in, the genome of the animal of interest, or, if endogenous target gene sequences are present, they may either be overexpressed or, alternatively, be disrupted in order to underexpress or inactivate target gene expression, such as described for the disruption of apoE in mice (Plump et al., 1992, Cell 71: 343-353).

In order to overexpress a target gene sequence, the coding portion of the target gene sequence may be ligated to a regulatory sequence which is capable of driving gene expression in the animal and cell type of interest. Such regulatory regions will be well known to those of skill in the art, and may be utilized in the absence of undue experimentation.

The use of such a genetically engineered animal-based system is described in detail in the example provided in Section 7, below, for the target gene bcl-2 (see Table 1, in Section 5.4.1, above). Briefly, bcl-2 expression first

was shown to be down-regulated in the apoE-deficient atherosclerosis mouse model. Then, a transgenic mouse was engineered bearing the human bcl-2 gene under the control of a promoter which is induced under atherogenic conditions. To test the effect of the induction of bcl-2 under such conditions, the transgenic mouse is crossed with the apoE-deficient mouse. apoE-deficient progeny bearing the highly expressible bcl-2 gene are then examined for plaque formation and development. Reduction in plaque formation and development in these progeny confirms the effectiveness of intervening in cardiovascular disease through this target gene.

For underexpression of an endogenous target gene sequence, such a sequence may be isolated and engineered such that when reintroduced into the genome of the animal of interest, the endogenous target gene alleles will be inactivated. Preferably, the engineered target gene sequence is introduced via gene targeting such that the endogenous target sequence is disrupted upon integration of the engineered target gene sequence into the animal's genome. Gene targeting is discussed, below, in this Section.

Animals of any species, including, but not limited to, mice, rats, rabbits, guinea pigs, pigs, micro-pigs, goats, and non-human primates, e.g., baboons, monkeys, and chimpanzees may be used to generate cardiovascular disease animal models.

Any technique known in the art may be used to introduce a target gene transgene into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to pronuclear microinjection (Hoppe, P.C. and Wagner, T.E., 1989, U.S. Pat. No. 4,873,191); retrovirus mediated gene transfer into germ lines (Van der Putten et al., 1985, Proc. Natl. Acad. Sci., USA 82:6148-6152); gene targeting in embryonic stem cells (Thompson et al., 1989, Cell 56:313-321); electroporation of embryos (Lo, 1983, Mol Cell. Biol. 3:1803-1814); and sperm-mediated gene transfer (Lavitrano et al., 1989, Cell 57:717-

723); etc. For a review of such techniques, see Gordon, 1989, Transgenic Animals, Intl. Rev. Cytol. 115:171-229, which is incorporated by reference herein in its entirety.

The present invention provides for transgenic animals that carry the transgene in all their cells, as well as animals which carry the transgene in some, but not all their cells, i.e., mosaic animals. The transgene may be integrated as a single transgene or in concatamers, e.g., head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular cell type by following, for example, the teaching of Lasko et al. (Lasko, M. et al., 1992, Proc. Natl. Acad. Sci. USA 89: 6232-6236). The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art. When it is desired that the target gene transgene be integrated into the chromosomal site of the endogenous target gene, gene targeting is preferred. Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to the endogenous target gene of interest are designed for the purpose of integrating, via homologous recombination with chromosomal sequences, into and disrupting the function of the nucleotide sequence of the endogenous target gene. The transgene may also be selectively introduced into a particular cell type, thus inactivating the endogenous gene of interest in only that cell type, by following, for example, the teaching of Gu et al. (Gu, et al., 1994, Science 265: 103-106). The regulatory sequences required for such a cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art. Recombinant methods for expressing target genes are described in Section 5.4.2, above.

Once transgenic animals have been generated, the expression of the recombinant target gene and protein may be assayed utilizing standard techniques. Initial screening may

be accomplished by Southern blot analysis or PCR techniques to analyze animal tissues to assay whether integration of the transgene has taken place. The level of mRNA expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques which include but are not limited to Northern blot analysis of tissue samples obtained from the animal, in situ hybridization analysis, and RT-PCR. Samples of target gene-expressing tissue, may also be evaluated immunocytochemically using antibodies specific for the target gene transgene gene product of interest.

The target gene transgenic animals that express target gene mRNA or target gene transgene peptide (detected immunocytochemically, using antibodies directed against the target gene product's epitopes) at easily detectable levels should then be further evaluated to identify those animals which display characteristic cardiovascular disease symptoms. Such symptoms may include, for example, increased prevalence and size of fatty streaks and/or cardiovascular disease plaques.

Additionally, specific cell types within the transgenic animals may be analyzed and assayed for cellular phenotypes characteristic of cardiovascular disease. In the case of monocytes, such phenotypes may include but are not limited to increases in rates of LDL uptake, adhesion to endothelial cells, transmigration, foam cell formation, fatty streak formation, and production of foam cell specific products. Cellular phenotype assays are discussed in detail in Section 5.4.4.2, below. Further, such cellular phenotypes may include a particular cell type's fingerprint pattern of expression as compared to known fingerprint expression profiles of the particular cell type in animals exhibiting cardiovascular disease symptoms. Fingerprint profiles are described in detail in Section 5.8.1, below. Such transgenic animals serve as suitable model systems for cardiovascular disease.

Once target gene transgenic founder animals are produced, they may be bred, inbred, outbred, or crossbred to

produce colonies of the particular animal. Examples of such breeding strategies include but are not limited to: outbreeding of founder animals with more than one integration site in order to establish separate lines; inbreeding of separate lines in order to produce compound target gene transgenics that express the target gene transgene of interest at higher levels because of the effects of additive expression of each target gene transgene; crossing of heterozygous transgenic animals to produce animals homozygous for a given integration site in order both to augment expression and eliminate the possible need for screening of animals by DNA analysis; crossing of separate homozygous lines to produce compound heterozygous or homozygous lines; breeding animals to different inbred genetic backgrounds so as to examine effects of modifying alleles on expression of the target gene transgene and the development of cardiovascular disease symptoms. One such approach is to cross the target gene transgenic founder animals with a wild type strain to produce an F1 generation that exhibits cardiovascular disease symptoms. The F1 generation may then be inbred in order to develop a homozygous line, if it is found that homozygous target gene transgenic animals are viable.

5.4.4.2. CELL-BASED ASSAYS

Cells that contain and express target gene sequences which encode target gene protein, and, further, exhibit cellular phenotypes associated with cardiovascular disease, may be utilized to identify compounds that exhibit anti-cardiovascular disease activity.

Such cells may include non-recombinant monocyte cell lines, such as U937 (ATCC# CRL-1593), THP-1 (ATCC# TIB-202), and P388D1 (ATCC# TIB-63); endothelial cells such as HUVEC's and bovine aortic endothelial cells (BAEC's); as well as generic mammalian cell lines such as HeLa cells and COS cells, e.g., COS-7 (ATCC# CRL-1651). Further, such cells may include recombinant, transgenic cell lines. For example, the

cardiovascular disease animal models of the invention, discussed, above, in Section 5.4.4.1, may be used to generate cell lines, containing one or more cell types involved in cardiovascular disease, that can be used as cell culture models for this disorder. While primary cultures derived from the cardiovascular disease transgenic animals of the invention may be utilized, the generation of continuous cell lines is preferred. For examples of techniques which may be used to derive a continuous cell line from the transgenic animals, see Small et al., 1985, Mol. Cell Biol. 5:642-648.

Alternatively, cells of a cell type known to be involved in cardiovascular disease may be transfected with sequences capable of increasing or decreasing the amount of target gene expression within the cell. For example, target gene sequences may be introduced into, and overexpressed in, the genome of the cell of interest, or, if endogenous target gene sequences are present, they may be either overexpressed or, alternatively disrupted in order to underexpress or inactivate target gene expression.

In order to overexpress a target gene sequence, the coding portion of the target gene sequence may be ligated to a regulatory sequence which is capable of driving gene expression in the cell type of interest. Such regulatory regions will be well known to those of skill in the art, and may be utilized in the absence of undue experimentation. Recombinant methods for expressing target genes are described in Section 5.4.2, above.

For underexpression of an endogenous target gene sequence, such a sequence may be isolated and engineered such that when reintroduced into the genome of the cell type of interest, the endogenous target gene alleles will be inactivated. Preferably, the engineered target gene sequence is introduced via gene targeting such that the endogenous target sequence is disrupted upon integration of the engineered target gene sequence into the cell's genome. Transfection of host cells with target genes is discussed, above, in Section 5.4.4.1.

Cells treated with compounds or transfected with target genes can be examined for phenotypes associated with cardiovascular disease. In the case of monocytes, such phenotypes include but are not limited to increases in rates of LDL uptake, adhesion to endothelial cells, transmigration, foam cell formation, fatty streak formation, and production by foam cells of growth factors such as bFGF, IGF-I, VEGF, IL-1, M-CSF, TGFβ, TGFα, TNFα, HB-EGF, PDGF, IFN-γ, and GM-CSF. Transmigration rates, for example, may be measured using the in vitro system of Navab et al., described in Section 5.1.1.3, above, by quantifying the number of monocytes that migrate across the endothelial monolayer and into the collagen layer of the subendothelial space.

Similarly, HUVEC's can be treated with test compounds or transfected with genetically engineered target genes described in Section 5.4.2, above. The HUVEC's can then be examined for phenotypes associated with cardiovascular disease, including, but not limited to changes in cellular morphology, cell proliferation, cell migration, and mononuclear cell adhesion; or for the effects on production of other proteins involved in cardiovascular disease such as ICAM, VCAM, PDGF-β, and E-selectin.

Transfection of target gene sequence nucleic acid may be accomplished by utilizing standard techniques. See, for example, Ausubel, 1989, *supra*. Transfected cells should be evaluated for the presence of the recombinant target gene sequences, for expression and accumulation of target gene mRNA, and for the presence of recombinant target gene protein production. In instances wherein a decrease in target gene expression is desired, standard techniques may be used to demonstrate whether a decrease in endogenous target gene expression and/or in target gene product production is achieved.

5.5. SCREENING ASSAYS FOR COMPOUNDS THAT INTERACT WITH THE TARGET GENE PRODUCT AND/OR MODULATE TARGET GENE EXPRESSION

The following assays are designed to identify compounds that bind to target gene products, bind to other cellular or extracellular proteins that interact with a target gene product, and interfere with the interaction of the target gene product with other cellular or extracellular proteins. Such compounds can act as the basis for amelioration of such cardiovascular diseases as atherosclerosis, ischemia/reperfusion, hypertension, restenosis, and arterial inflammation by modulating the activity of the protein products of target genes. Such compounds may also act as the basis for the amelioration of fibroproliferative and oncogenic related disorders, including tumorigenesis and the vascularization of tumors. Such compounds may include, but are not limited to peptides, antibodies, or small organic or inorganic compounds. Methods for the identification of such compounds are described in Section 5.5.1, below. Such compounds may also include other cellular proteins. Methods for the identification of such cellular proteins are described, below, in Section 5.5.2.

For example, such assays can be used to identify compounds that bind to the rchd534 protein, the rchd534-long protein, or the fchd540 protein, including compounds that interfere with the interaction of the the rchd534 protein with itself, the rchd534-long protein with itself, or the fchd540 protein with either the rchd534 protein or the rchd534-long protein.

Compounds identified via assays such as those described herein may be useful, for example, in elaborating the biological function of the target gene product, and for ameliorating cardiovascular disease. In instances whereby a cardiovascular disease condition results from an overall lower level of target gene expression and/or target gene product in a cell or tissue, compounds that interact with the target gene product may include compounds which accentuate or amplify the activity of the bound target gene protein. Such

compounds would bring about an effective increase in the level of target gene product activity, thus ameliorating symptoms.

In some cases, a target gene observed to be up-regulated under disease conditions may be exerting a protective effect. Compounds that enhance the expression of such up-regulated genes, or the activity of their gene products, would also ameliorate disease symptoms, especially in individuals whose target gene is not normally up-regulated.

In other instances mutations within the target gene may cause aberrant types or excessive amounts of target gene proteins to be made which have a deleterious effect that leads to cardiovascular disease. Similarly, physiological conditions may cause an excessive increase in target gene expression leading to cardiovascular disease. In such cases, compounds that bind target gene protein may be identified that inhibit the activity of the bound target gene protein. Assays for testing the effectiveness of compounds, identified by, for example, techniques such as those described in this Section are discussed, below, in Section 5.5.4.

5.5.1. IN VITRO SCREENING ASSAYS FOR COMPOUNDS THAT BIND TO THE TARGET GENE PRODUCT

In vitro systems may be designed to identify compounds capable of binding the target gene of the invention. Such compounds may include, but are not limited to, peptides made of D-and/or L-configuration amino acids (in, for example, the form of random peptide libraries; see e.g., Lam, K.S. et al., 1991, Nature 354:82-84), phosphopeptides (in, for example, the form of random or partially degenerate, directed phosphopeptide libraries; see, e.g., Songyang, Z. et al., 1993, Cell 72:767-778), antibodies, and small organic or inorganic molecules. Compounds identified may be useful, for example, in modulating the activity of target gene proteins, preferably mutant target gene proteins, may be useful in elaborating the

biological function of the target gene protein, may be utilized in screens for identifying compounds that disrupt normal target gene interactions, or may in themselves disrupt such interactions.

The principle of the assays used to identify compounds that bind to the target gene protein involves preparing a reaction mixture of the target gene protein and the test compound under conditions and for a time sufficient to allow the two components to interact and bind, thus forming a complex which can be removed and/or detected in the reaction mixture. These assays can be conducted in a variety of ways. For example, one method to conduct such an assay would involve anchoring the target gene or the test substance onto a solid phase and detecting target gene/test substance complexes anchored on the solid phase at the end of the reaction. In one embodiment of such a method, the target gene protein may be anchored onto a solid surface, and the test compound, which is not anchored, may be labeled, either directly or indirectly.

In practice, microtitre plates are conveniently utilized. The anchored component may be immobilized by non-covalent or covalent attachments. Non-covalent attachment may be accomplished simply by coating the solid surface with a solution of the protein and drying. Alternatively, an immobilized antibody, preferably a monoclonal antibody, specific for the protein may be used to anchor the protein to the solid surface. The surfaces may be prepared in advance and stored.

In order to conduct the assay, the nonimmobilized component is added to the coated surface containing the anchored component. After the reaction is complete, unreacted components are removed (e.g., by washing) under conditions such that any complexes formed will remain immobilized on the solid surface. The detection of complexes anchored on the solid surface can be accomplished in a number of ways. Where the previously nonimmobilized component is pre-labeled, the detection of label immobilized on the

surface indicates that complexes were formed. Where the previously nonimmobilized component is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface; e.g., using a labeled antibody specific for the previously nonimmobilized component (the antibody, in turn, may be directly labeled or indirectly labeled with a labeled anti-Ig antibody).

Alternatively, a reaction can be conducted in a liquid phase, the reaction products separated from unreacted components, and complexes detected; e.g., using an immobilized antibody specific for target gene product or the test compound to anchor any complexes formed in solution, and a labeled antibody specific for the other component of the possible complex to detect anchored complexes.

Compounds that are shown to bind to a particular target gene product through one of the methods described above can be further tested for their ability to elicit a biochemical response from the target gene protein. A particular embodiment is described herein for receptor proteins involved in signal transduction, including but not limited to the rchd523 gene product. Compounds that interact with a target gene product receptor domain, can be screened for their ability to function as ligands, i.e., to bind to the receptor protein in a manner that triggers the signal transduction pathway. Useful receptor fragments or analogs in the invention are those which interact with ligand. The receptor component can be assayed functionally, i.e., for its ability to bind ligand and mobilize Ca^{++} (see below). These assays include, as components, ligand and a recombinant target gene product (or a suitable fragment or analog) configured to permit detection of binding.

For example, and not by way of limitation, a recombinant receptor may be used to screen for ligands by its ability to mediate ligand-dependent mobilization of calcium. Cells, preferably myeloma cells or Xenopus oocytes, transfected with a target gene expression vector (constructed according to the methods described in Section 5.4.2, above)

are loaded with FURA-2 or INDO-1 by standard techniques. Mobilization of Ca^{2+} induced by ligand is measured by fluorescence spectroscopy as previously described (Grynkiewicz et al., 1985, *J. Biol. Chem.* 260:3440). Ligands that react with the target gene product receptor domain, therefore, can be identified by their ability to produce a fluorescent signal. Their receptor binding activities can be quantified and compared by measuring the level of fluorescence produced over background.

The rchd523 gene product consists of a G protein-coupled receptor with multiple transmembrane domains. The Ca^{2+} mobilization assay, therefore, can be used to screen compounds that are ligands of the rchd523 receptor. This screening method is described in detail with respect to rchd523 in the example in Section 12, below. Identification of rchd523 ligand, and measuring the activity of the ligand-receptor complex, leads to the identification of antagonists of this interaction, as described in Section 5.5.3, below. Such antagonists are useful in the treatment of cardiovascular disease.

5.5.2. ASSAYS FOR CELLULAR OR EXTRACELLULAR PROTEINS THAT INTERACT WITH THE TARGET GENE PRODUCT

Any method suitable for detecting protein-protein interactions may be employed for identifying novel target protein-cellular or extracellular protein interactions. These methods are outlined in Section 5.2., *supra*, for the identification of pathway genes, and may be utilized herein with respect to the identification of proteins which interact with identified target proteins. In such a case, the target gene serves as the known "bait" gene.

5.5.3. ASSAYS FOR COMPOUNDS THAT INTERFERE WITH INTERACTION BETWEEN TARGET GENE PRODUCT AND OTHER COMPOUNDS

The target gene proteins of the invention may, in vivo, interact with one or more cellular or extracellular

proteins. Such proteins may include, but are not limited to, those proteins identified via methods such as those described, above, in Section 5.5.2. For the purposes of this discussion, target gene products and such cellular and extracellular proteins are referred to herein as "binding partners". Compounds that disrupt such interactions may be useful in regulating the activity of the target gene proteins, especially mutant target gene proteins. Such compounds may include, but are not limited to molecules such as antibodies, peptides, and the like described in Section 5.5.1. above.

The basic principle of the assay systems used to identify compounds that interfere with the interaction between the target gene protein, and its cellular or extracellular protein binding partner or partners involves preparing a reaction mixture containing the target gene protein and the binding partner under conditions and for a time sufficient to allow the two proteins to interact and bind, thus forming a complex. In order to test a compound for inhibitory activity, the reaction mixture is prepared in the presence and absence of the test compound. The test compound may be initially included in the reaction mixture or may be added at a time subsequent to the addition of target gene and its cellular or extracellular binding partner. Control reaction mixtures are incubated without the test compound or with a placebo. The formation of any complexes between the target gene protein and the cellular or extracellular binding partner is then detected. The formation of a complex in the control reaction, but not in the reaction mixture containing the test compound, indicates that the compound interferes with the interaction of the target gene protein and the interactive binding partner protein. Additionally, complex formation within reaction mixtures containing the test compound and a normal target gene protein may also be compared to complex formation within reaction mixtures containing the test compound and mutant target gene protein. This comparison may be important in

those cases wherein it is desirable to identify compounds that disrupt interactions of mutant but not normal target gene proteins.

The assay for compounds that interfere with the interaction of the binding partners can be conducted in a heterogeneous or homogeneous format. Heterogeneous assays involve anchoring one of the binding partners onto a solid phase and detecting complexes anchored on the solid phase at the end of the reaction. In homogeneous assays, the entire reaction is carried out in a liquid phase. In either approach, the order of addition of reactants can be varied to obtain different information about the compounds being tested. For example, test compounds that interfere with the interaction between the binding partners, e.g., by competition, can be identified by conducting the reaction in the presence of the test substance; i.e., by adding the test substance to the reaction mixture prior to or simultaneously with the target gene protein and interactive cellular or extracellular protein. Alternatively, test compounds that disrupt preformed complexes, e.g. compounds with higher binding constants that displace one of the binding partners from the complex, can be tested by adding the test compound to the reaction mixture after complexes have been formed. The various formats are described briefly below.

In a heterogeneous assay system, either the target gene protein or the interactive cellular or extracellular binding partner protein, is anchored onto a solid surface, and its binding partner, which is not anchored, is labeled, either directly or indirectly. In practice, microtitre plates are conveniently utilized. The anchored species may be immobilized by non-covalent or covalent attachments. Non-covalent attachment may be accomplished simply by coating the solid surface with a solution of the protein and drying. Alternatively, an immobilized antibody specific for the protein may be used to anchor the protein to the solid surface. The surfaces may be prepared in advance and stored.

In order to conduct the assay, the binding partner of the immobilized species is exposed to the coated surface with or without the test compound. After the reaction is complete, unreacted components are removed (e.g., by washing) and any complexes formed will remain immobilized on the solid surface. The detection of complexes anchored on the solid surface can be accomplished in a number of ways. Where the binding partner was pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the binding partner is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface; e.g., using a labeled antibody specific for the binding partner (the antibody, in turn, may be directly labeled or indirectly labeled with a labeled anti-Ig antibody). Depending upon the order of addition of reaction components, test compounds which inhibit complex formation or which disrupt preformed complexes can be detected.

Alternatively, the reaction can be conducted in a liquid phase in the presence or absence of the test compound, the reaction products separated from unreacted components, and complexes detected; e.g., using an immobilized antibody specific for one binding partner to anchor any complexes formed in solution, and a labeled antibody specific for the other binding partner to detect anchored complexes. Again, depending upon the order of addition of reactants to the liquid phase, test compounds which inhibit complex or which disrupt preformed complexes can be identified.

In an alternate embodiment of the invention, a homogeneous assay can be used. In this approach, a preformed complex of the target gene protein and the interactive cellular or extracellular protein is prepared in which one of the binding partners is labeled, but the signal generated by the label is quenched due to complex formation (see, e.g., U.S. Patent No. 4,109,496 by Rubenstein which utilizes this approach for immunoassays). The addition of a test substance that competes with and displaces one of the binding partners from the preformed complex will result in the generation of a

signal above background. In this way, test substances which disrupt target gene protein-cellular or extracellular protein interaction can be identified.

In a particular embodiment, the target gene protein can be prepared for immobilization using recombinant DNA techniques described in Section 5.4.2, *supra*. For example, the target gene coding region can be fused to a glutathione-S-transferase (GST) gene, using a fusion vector such as pGEX-5X-1, in such a manner that its binding activity is maintained in the resulting fusion protein. The interactive cellular or extracellular protein can be purified and used to raise a monoclonal antibody, using methods routinely practiced in the art and described above, in Section 5.4.3. This antibody can be labeled with the radioactive isotope ^{125}I, for example, by methods routinely practiced in the art. In a heterogeneous assay, e.g., the GST-target gene fusion protein can be anchored to glutathione-agarose beads. The interactive cellular or extracellular binding partner protein can then be added in the presence or absence of the test compound in a manner that allows interaction and binding to occur. At the end of the reaction period, unbound material can be washed away, and the labeled monoclonal antibody can be added to the system and allowed to bind to the complexed binding partners. The interaction between the target gene protein and the interactive cellular or extracellular binding partner protein can be detected by measuring the amount of radioactivity that remains associated with the glutathione-agarose beads. A successful inhibition of the interaction by the test compound will result in a decrease in measured radioactivity.

Alternatively, the GST-target gene fusion protein and the interactive cellular or extracellular binding partner protein can be mixed together in liquid in the absence of the solid glutathione-agarose beads. The test compound can be added either during or after the binding partners are allowed to interact. This mixture can then be added to the glutathione-agarose beads and unbound material is washed

away. Again the extent of inhibition of the binding partner interaction can be detected by adding the labeled antibody and measuring the radioactivity associated with the beads.

In another embodiment of the invention, these same techniques can be employed using peptide fragments that correspond to the binding domains of the target gene protein and the interactive cellular or extracellular protein, respectively, in place of one or both of the full length proteins. Any number of methods routinely practiced in the art can be used to identify and isolate the protein's binding site. These methods include, but are not limited to, mutagenesis of one of the genes encoding the proteins and screening for disruption of binding in a co-immunoprecipitation assay. Compensating mutations in the target gene can be selected. Sequence analysis of the genes encoding the respective proteins will reveal the mutations that correspond to the region of the protein involved in interactive binding. Alternatively, one protein can be anchored to a solid surface using methods described in this Section above, and allowed to interact with and bind to its labeled binding partner, which has been treated with a proteolytic enzyme, such as trypsin. After washing, a short, labeled peptide comprising the binding domain may remain associated with the solid material, which can be isolated and identified by amino acid sequencing. Also, once the gene coding for the for the cellular or extracellular protein is obtained, short gene segments can be engineered to express peptide fragments of the protein, which can then be tested for binding activity and purified or synthesized.

For example, and not by way of limitation, target gene can be anchored to a solid material as described above in this Section by making a GST-target gene fusion protein and allowing it to bind to glutathione agarose beads. The interactive cellular or extracellular binding partner protein can be labeled with a radioactive isotope, such as ^{35}S, and cleaved with a proteolytic enzyme such as trypsin. Cleavage products can then be added to the anchored GST-target gene

fusion protein and allowed to bind. After washing away unbound peptides, labeled bound material, representing the cellular or extracellular binding partner protein binding domain, can be eluted, purified, and analyzed for amino acid sequence by techniques well known in the art; e.g., using the Edman degradation procedure (see e.g., Creighton, 1983, Proteins: Structures and Molecular Principles, W.H. Freeman & Co., N.Y., pp. 34-49). Peptides so identified can be produced, using techniques well known in the art, either synthetically (see e.g., Creighton, 1983, *supra* at pp. 50-60) or, if the gene has already been isolated, by using recombinant DNA technology, as described in Section 5.4.2, *supra*.

A particular embodiment of the invention features a method of screening candidate compounds for their ability to antagonize the interaction between ligand and the receptor domain of a target gene product, including but not limited to the receptor domain of the rchd523 gene product. The rchd523 gene product, which is a G protein-coupled receptor protein containing multiple transmembrane domains, is especially useful in screening for antagonists of ligand-receptor interactions. The method involves: a) mixing a candidate antagonist compound with a first compound which includes a recombinant target gene product comprising a receptor domain (or ligand-binding fragment or analog) on the one hand and with a second compound which includes ligand on the other hand; b) determining whether the first and second compounds bind; and c) identifying antagonistic compounds as those which interfere with the binding of the first compound to the second compound and/or which reduce the ligand-mediated release of intracellular Ca^{++}.

By an "antagonist" is meant a molecule which inhibits a particular activity, in this case, the ability of ligand to interact with a target gene product receptor domain and/or to trigger the biological events resulting from such an interaction (e.g., release of intracellular Ca^{++}). Preferred therapeutics include antagonists, e.g., peptide

fragments (particularly, fragments derived from the N-terminal extracellular domain), antibodies (particularly, antibodies which recognize and bind the N-terminal extracellular domain), or drugs, which block ligand or target gene product function by interfering with the ligand-receptor interaction.

Because the receptor component of the target gene product can be produced by recombinant techniques and because candidate antagonists may be screened in vitro, the instant invention provides a simple and rapid approach to the identification of useful therapeutics.

Specific receptor fragments of interest include any portions of the target gene products that are capable of interaction with ligand, for example, all or part of the N-terminal extracellular domain. Such portions include the transmembrane segments and portions of the receptor deduced to be extracellular. Such fragments may be useful as antagonists (as described above), and are also useful as immunogens for producing antibodies which neutralize the activity of the target gene product in vivo (e.g., by interfering with the interaction between the receptor and ligand; see below). Extracellular regions may be identified by comparison with related proteins of similar structure (e.g., other members of the G-protein-coupled receptor superfamily); useful regions are those exhibiting homology to the extracellular domains of well-characterized members of the family.

Alternatively, from the primary amino acid sequence, the secondary protein structure and, therefore, the extracellular domain regions may be deduced semi-empirically using a hydrophobicity/hydrophilicity calculation such as the Chou-Fasman method (see, e.g., Chou and Fasman, *Ann. Rev. Biochem.* 47:251, 1978). Hydrophilic domains, particularly ones surrounded by hydrophobic stretches (e.g., transmembrane domains) present themselves as strong candidates for extracellular domains. Finally, extracellular domains may be

identified experimentally using standard enzymatic digest analysis, e.g., tryptic digest analysis.

Candidate fragments (e.g., all or part of the transmembrane segments or any extracellular fragment) are tested for interaction with ligand by the assays described herein (e.g., the assay described above). Such fragments are also tested for their ability to antagonize the interaction between ligand and its endogenous receptor using the assays described herein. Analogs of useful receptor fragments (as described above) may also be produced and tested for efficacy as screening components or antagonists (using the assays described herein); such analogs are also considered to be useful in the invention.

Of particular interest are receptor fragments encompassing the extracellular main-terminal domain (or a ligand binding fragment thereof). Also of interest are the target gene product extracellular loops. Peptide fragments derived from these extracellular loops may also be used as antagonists, particularly if the loops cooperate with the amino-terminal domain to facilitate ligand binding. Alternatively, such loops and extracellular N-terminal domain (as well as the full length target gene product) provide immunogens for producing anti-target gene product antibodies.

Binding of ligand to its receptor may be assayed by any of the methods described above in Section 5.5.1. Preferably, cells expressing recombinant target gene product (or a suitable target gene product fragment or analog) are immobilized on a solid substrate (e.g., the wall of a microtitre plate or a column) and reacted with detectably-labelled ligand (as described above). Binding is assayed by the detection label in association with the receptor component (and, therefore, in association with the solid substrate). Binding of labelled ligand to receptor-bearing cells is used as a "control" against which antagonist assays are measured. The antagonist assays involve incubation of the target gene product-bearing cells with an appropriate amount of candidate antagonist. To this mix, an equivalent

amount to labelled ligand is added. An antagonist useful in the invention specifically interferes with labelled ligand binding to the immobilized receptor-expressing cells.

An antagonist is then tested for its ability to interfere with ligand function, i.e., to specifically interfere with labelled ligand binding without resulting in signal transduction normally mediated by the receptor. To test this using a functional assay, stably transfected cell lines containing the target gene product can be produced as described herein and reporter compounds such as the calcium binding agent, FURA-2, loaded into the cytoplasm by standard techniques. Stimulation of the heterologous target gene product with ligand or another agonist leads to intracellular calcium release and the concomitant fluorescence of the calcium-FURA-2 complex. This provides a convenient means for measuring agonist activity. Inclusion of potential antagonists along with ligand allows for the screening and identification of authentic receptor antagonists as those which effectively block ligand binding without producing fluorescence (i.e., without causing the mobilization of intracellular Ca^{++}). Such an antagonist may be expected to be a useful therapeutic agent for cardiovascular disorders.

Appropriate candidate antagonists include target gene product fragments, particularly fragments containing a ligand-binding portion adjacent to or including one or more transmembrane segments or an extracellular domain of the receptor (described above); such fragments would preferably including five or more amino acids. Other candidate antagonists include analogs of ligand and other peptides as well as non-peptide compounds and anti-target gene product antibodies designed or derived from analysis of the receptor.

This screening method is described in detail with respect to the rchd523 gene in the example in Section 12, below. Because the rchd523 gene product is a G protein-coupled receptor, antagonists of the interaction between the rchd523 gene product and its natural ligand provide excellent

candidates for compounds effective in the treatment of cardiovascular disease.

5.5.4. ASSAYS FOR AMELIORATION OF CARDIOVASCULAR DISEASE SYMPTOMS

Any of the binding compounds, including but not limited to compounds such as those identified in the foregoing assay systems, may be tested for the ability to ameliorate cardiovascular disease symptoms. Cell-based and animal model-based assays for the identification of compounds exhibiting such an ability to ameliorate cardiovascular disease symptoms are described below.

First, cell-based systems such as those described, above, in Section 5.4.4.2., may be used to identify compounds which may act to ameliorate cardiovascular disease symptoms. For example, such cell systems may be exposed to a compound, suspected of exhibiting an ability to ameliorate cardiovascular disease symptoms, at a sufficient concentration and for a time sufficient to elicit such an amelioration of cardiovascular disease symptoms in the exposed cells. After exposure, the cells are examined to determine whether one or more of the cardiovascular disease cellular phenotypes has been altered to resemble a more normal or more wild type, non-cardiovascular disease phenotype. For example, and not by way of limitation, in the case of monocytes, such more normal phenotypes may include but are not limited to decreased rates of LDL uptake, adhesion to endothelial cells, transmigration, foam cell formation, fatty streak formation, and production by foam cells of growth factors such as bFGF, IGF-I, VEGF, IL-1, M-CSF, TGFβ, TGFα, TNFα, HB-EGF, PDGF, IFN-γ, and GM-CSF. Transmigration rates, for example, may be measured using the in vitro system of Navab et al., described in Section 5.1.1.3, above, by quantifying the number of monocytes that migrate across the endothelial monolayer and into the collagen layer of the subendothelial space.

In addition, animal-based cardiovascular disease systems, such as those described, above, in Section 5.4.4.1, may be used to identify compounds capable of ameliorating cardiovascular disease symptoms. Such animal models may be used as test substrates for the identification of drugs, pharmaceuticals, therapies, and interventions which may be effective in treating cardiovascular disease. For example, animal models may be exposed to a compound, suspected of exhibiting an ability to ameliorate cardiovascular disease symptoms, at a sufficient concentration and for a time sufficient to elicit such an amelioration of cardiovascular disease symptoms in the exposed animals. The response of the animals to the exposure may be monitored by assessing the reversal of disorders associated with cardiovascular disease, for example, by counting the number of atherosclerotic plaques and/or measuring their size before and after treatment.

Further, both cell-based systems and animal-based systems as described herein may be used to identify compounds which act to ameliorate symptoms of fibroproliferative and oncogenic related disorders, including tumorigenesis and the vascularization of tumors.

With regard to intervention, any treatments which reverse any aspect of symptoms of cardiovascular disease or fibroproliferative and oncogenic related disorders should be considered as candidates for human therapeutic intervention. Dosages of test agents may be determined by deriving dose-response curves, as discussed in Section 5.7.1, below.

Additionally, gene expression patterns may be utilized to assess the ability of a compound to ameliorate symptoms of cardiovascular disease or fibroproliferative and oncogenic related disorders. For example, the expression pattern of one or more fingerprint genes may form part of a "fingerprint profile" which may be then be used in such an assessment. "Fingerprint profile", as used herein, refers to the pattern of mRNA expression obtained for a given tissue or cell type under a given set of conditions. Such conditions

may include, but are not limited to, atherosclerosis, ischemia/reperfusion, hypertension, restenosis, and arterial inflammation, including any of the control or experimental conditions described in the paradigms of Section 5.1.1, above. Fingerprint profiles may be generated, for example, by utilizing a differential display procedure, as discussed, above, in Section 5.1.2, Northern analysis and/or RT-PCR. Any of the gene sequences described, above, in Section 5.4.1. may be used as probes and/or PCR primers for the generation and corroboration of such fingerprint profiles.

Fingerprint profiles may be characterized for known states, either cardiovascular disease or normal, within the cell- and/or animal-based model systems. Subsequently, these known fingerprint profiles may be compared to ascertain the effect a test compound has to modify such fingerprint profiles, and to cause the profile to more closely resemble that of a more desirable fingerprint.

For example, administration of a compound may cause the fingerprint profile of a cardiovascular disease model system to more closely resemble the control system. Administration of a compound may, alternatively, cause the fingerprint profile of a control system to begin to mimic a cardiovascular disease state. Such a compound may, for example, be used in further characterizing the compound of interest, or may be used in the generation of additional animal models.

5.5.5. MONITORING OF EFFECTS DURING CLINICAL TRIALS

Monitoring the influence of compounds on cardiovascular disease states may be applied not only in basic drug screening, but also in clinical trials. In such clinical trials, the expression of a panel of genes that have been discovered in one of the paradigms described in Section 5.1.1.1 through 5.1.1.6 may be used as a "read out" of a particular drug's effect on a cardiovascular disease state.

For example, and not by way of limitation, Paradigm A provides for the identification of fingerprint genes that are up-regulated in monocytes treated with oxidized LDL. Thus, to study the effect of anti-oxidant drugs, for example, in a clinical trial, blood may be drawn from patients before and at different stages during treatment with such a drug. Their monocytes may then be isolated and RNA prepared and analyzed by differential display as described in Sections 6.1.1 and 6.1.2. The levels of expression of these fingerprint genes may be quantified by Northern blot analysis or RT-PCR, as described in Section 6.1.2, or by one of the methods described in Section 5.8.1, or alternatively by measuring the amount of protein produced, by one of the methods described in Section 5.8.2. In this way, the fingerprint profiles may serve as surrogate markers indicative of the physiological response of monocytes that have taken up oxidized LDL. Accordingly, this response state may be determined before, and at various points during, drug treatment. This method is described in further detail in the example in Section 10, below.

This method may also be applied to the other paradigms disclosed herein. For example, and not by way of limitation, the fingerprint profile of Paradigm B reveals that bcl-2 and glutathione peroxidase are both down-regulated in the monocytes of patients exposed to a high lipid diet, e.g. cholesterol or fat, that leads to high serum LDL levels. Drugs may be tested, for example, for their ability to ameliorate the effects of hypercholesterolemia in clinical trials. Patients with high LDL levels may have their monocytes isolated before, and at different stages after, drug treatment. The drug's efficacy may be measured by determining the degree of restored expression of bcl-2 and glutathione peroxidase, as described above for the Paradigm A fingerprint profile.

5.5.6. ASSAYS FOR COMPOUNDS THAT MODULATE EXPRESSION OF TARGET GENES

Compounds and other substances that modulate expression of target genes can be screened using in vitro cellular systems. In a manner analogous to the monitoring of compounds clinical samples described in Section 5.5.5, above, a sample of cells, such as a tissue culture is exposed to a test substance. Appropriate tissue culture cells include, but are not limited to, human umbilical vein endothelial cells (HUVECs), bovine aortic endothelial cells (BAECs), and 293 cells (embryonic human kidney cells). The RNA is then extracted from the cells. The level of transcription of a specific target gene can be detected using, for example, standard RT-PCR amplification techniques and/or Northern analysis (as described in the example in Section 6.1.2, below). Alternatively, the level of target protein production can be assayed by using antibodies that detect the target gene protein, as described in Section 5.8.2, below. The level of expression is compared to a control cell sample which was not exposed to the test substance.

Compounds that can be screened for modulation of expression of the target gene include, but are not limited to, small inorganic or organic molecules, peptides, such as peptide hormones analogs, steroid hormones, analogs of such hormones, and other proteins. Compounds that down-regulate expression include, but are not limited to, oligonucleotides that are complementary to the 5'-end of the mRNA of the target gene and inhibit transcription by forming triple helix structures, and ribozymes or antisense molecules which inhibit translation of the target gene mRNA. Techniques and strategies for designing such down-regulating test compounds are described in detail in Section 5.6, below.

5.6. COMPOUNDS AND METHODS FOR TREATMENT OF CARDIOVASCULAR DISEASE

Described below are methods and compositions whereby cardiovascular disease symptoms may be ameliorated. The methods and compositions described below may also be

applied to the amelioration of symptoms associated with fibroproliferative and oncogenic disorders, but by way of example, and not by limitation, are described in the subsections below in terms of cardiovascular disease.

Certain cardiovascular diseases are brought about, at least in part, by an excessive level of gene product, or by the presence of a gene product exhibiting an abnormal or excessive activity. As such, the reduction in the level and/or activity of such gene products would bring about the amelioration of cardiovascular disease symptoms. Techniques for the reduction of target gene expression levels or target gene product activity levels are discussed in Section 5.6.1, below.

Alternatively, certain other cardiovascular diseases are brought about, at least in part, by the absence or reduction of the level of gene expression, or a reduction in the level of a gene product's activity. As such, an increase in the level of gene expression and/or the activity of such gene products would bring about the amelioration of cardiovascular disease symptoms.

In some cases, the up-regulation of a gene in a disease state reflects a protective role for that gene product in responding to the disease condition. Enhancement of such a target gene's expression, or the activity of the target gene product, will reinforce the protective effect it exerts. Some cardiovascular disease states may result from an abnormally low level of activity of such a protective gene. In these cases also, an increase in the level of gene expression and/or the activity of such gene products would bring about the amelioration of cardiovascular disease symptoms. Techniques for increasing target gene expression levels or target gene product activity levels are discussed in Section 5.6.2, below.

5.6.1. COMPOUNDS THAT INHIBIT EXPRESSION, SYNTHESIS OR ACTIVITY OF MUTANT TARGET GENE ACTIVITY

As discussed above, target genes involved in cardiovascular disease disorders can cause such disorders via an increased level of target gene activity. As summarized in Table 1, above, and detailed in the examples in Sections 8 and 9, below, a number of genes are now known to be up-regulated in endothelial cells under disease conditions. Specifically, rchd005, rchd024, rchd032, and rchd036 are all up-regulated in endothelial cells treated with IL-1. Furthermore, rchd502, rchd523, rchd528, rchd534, COX II, and MnSOD are all up-regulated in endothelial cells subjected to shear stress. In some cases, such up-regulation may have a causative or exacerbating effect on the disease state. A variety of techniques may be utilized to inhibit the expression, synthesis, or activity of such target genes and/or proteins.

For example, compounds such as those identified through assays described, above, in Section 5.5, which exhibit inhibitory activity, may be used in accordance with the invention to ameliorate cardiovascular disease symptoms. As discussed in Section 5.5, above, such molecules may include, but are not limited to small organic molecules, peptides, antibodies, and the like. Inhibitory antibody techniques are described, below, in Section 5.6.1.2.

For example, compounds can be administered that compete with endogenous ligand for the rchd523 gene product. The resulting reduction in the amount of ligand-bound rchd523 gene transmembrane protein will modulated endothelial cell physiology. Compounds that can be particularly useful for this purpose include, for example, soluble proteins or peptides, such as peptides comprising one or more of the extracellular domains, or portions and/or analogs thereof, of the rchd523 gene product, including, for example, soluble fusion proteins such as Ig-tailed fusion proteins. (For a discussion of the production of Ig-tailed fusion proteins, see, for example, U.S. Patent No. 5,116,964.).

Alternatively, compounds, such as ligand analogs or antibodies, that bind to the rchd523 gene product receptor site, but do not activate the protein, (*e.g.*, receptor-ligand antagonists) can be effective in inhibiting rchd523 gene product activity.

Further, antisense and ribozyme molecules which inhibit expression of the target gene may also be used in accordance with the invention to inhibit the aberrant target gene activity. Such techniques are described, below, in Section 5.6.1.1. Still further, also as described, below, in Section 5.6.1.1, triple helix molecules may be utilized in inhibiting the aberrant target gene activity.

5.6.1.1. INHIBITORY ANTISENSE, RIBOZYME, TRIPLE HELIX, AND GENE INACTIVATION APPROACHES

Among the compounds which may exhibit the ability to ameliorate cardiovascular disease symptoms are antisense, ribozyme, and triple helix molecules. Such molecules may be designed to reduce or inhibit mutant target gene activity. Techniques for the production and use of such molecules are well known to those of skill in the art.

Antisense RNA and DNA molecules act to directly block the translation of mRNA by hybridizing to targeted mRNA and preventing protein translation.

Antisense approaches involve the design of oligonucleotides (either DNA or RNA) that are complementary to target gene mRNA. The antisense oligonucleotides will bind to the complementary target gene mRNA transcripts and prevent translation. Absolute complementarity, although preferred, is not required. A sequence "complementary" to a portion of an RNA, as referred to herein, means a sequence having sufficient complementarity to be able to hybridize with the RNA, forming a stable duplex; in the case of double-stranded antisense nucleic acids, a single strand of the duplex DNA may thus be tested, or triplex formation may be assayed. The ability to hybridize will depend on both the degree of complementarity and the length of the antisense

nucleic acid. Generally, the longer the hybridizing nucleic acid, the more base mismatches with an RNA it may contain and still form a stable duplex (or triplex, as the case may be). One skilled in the art can ascertain a tolerable degree of mismatch by use of standard procedures to determine the melting point of the hybridized complex.

Oligonucleotides that are complementary to the 5′ end of the message, e.g., the 5′ untranslated sequence up to and including the AUG initiation codon, should work most efficiently at inhibiting translation. However, sequences complementary to the 3′ untranslated sequences of mRNAs have recently shown to be effective at inhibiting translation of mRNAs as well. See generally, Wagner, R., 1994, Nature 372:333-335. Thus, oligonucleotides complementary to either the 5′- or 3′- non- translated, non-coding regions of the target gene could be used in an antisense approach to inhibit translation of endogenous target gene mRNA. Oligonucleotides complementary to the 5′ untranslated region of the mRNA should include the complement of the AUG start codon. Antisense oligonucleotides complementary to mRNA coding regions are less efficient inhibitors of translatisAon but could be used in accordance with the invention. Whether designed to hybridize to the 5′-, 3′- or coding region of target gene mRNA, antisense nucleic acids should be at least six nucleotides in length, and are preferably oligonucleotides ranging from 6 to about 50 nucleotides in length. In specific aspects the oligonucleotide is at least 10 nucleotides, at least 17 nucleotides, at least 25 nucleotides or at least 50 nucleotides.

Regardless of the choice of target sequence, it is preferred that *in vitro* studies are first performed to quantitate the ability of the antisense oligonucleotide to inhibit gene expression. It is preferred that these studies utilize controls that distinguish between antisense gene inhibition and nonspecific biological effects of oligonucleotides. It is also preferred that these studies compare levels of the target RNA or protein with that of an

internal control RNA or protein. Additionally, it is envisioned that results obtained using the antisense oligonucleotide are compared with those obtained using a control oligonucleotide. It is preferred that the control oligonucleotide is of approximately the same length as the test oligonucleotide and that the nucleotide sequence of the oligonucleotide differs from the antisense sequence no more than is necessary to prevent specific hybridization to the target sequence.

The oligonucleotides can be DNA or RNA or chimeric mixtures or derivatives or modified versions thereof, single-stranded or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone, for example, to improve stability of the molecule, hybridization, etc. The oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors in vivo), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger et al., 1989, Proc. Natl. Acad. Sci. U.S.A. 86:6553-6556; Lemaitre et al., 1987, Proc. Natl. Acad. Sci. 84:648-652; PCT Publication No. WO88/09810, published December 15, 1988) or the blood-brain barrier (see, *e.g.*, PCT Publication No. WO89/10134, published April 25, 1988), hybridization-triggered cleavage agents. (See, *e.g.*, Krol et al., 1988, BioTechniques 6:958-976) or intercalating agents. (See, *e.g.*, Zon, 1988, Pharm. Res. 5:539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

The antisense oligonucleotide may comprise at least one modified base moiety which is selected from the group including but not limited to 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine,

1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

The antisense oligonucleotide may also comprise at least one modified sugar moiety selected from the group including but not limited to arabinose, 2-fluoroarabinose, xylulose, and hexose.

In yet another embodiment, the antisense oligonucleotide comprises at least one modified phosphate backbone selected from the group consisting of a phosphorothioate, a phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal or analog thereof.

In yet another embodiment, the antisense oligonucleotide is an α-anomeric oligonucleotide. An α-anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other (Gautier et al., 1987, Nucl. Acids Res. 15:6625-6641). The oligonucleotide is a 2'-0-methylribonucleotide (Inoue et al., 1987, Nucl. Acids Res. 15:6131-6148), or a chimeric RNA-DNA analogue (Inoue et al., 1987, FEBS Lett. 215:327-330).

Oligonucleotides of the invention may be synthesized by standard methods known in the art, *e.g.* by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples, phosphorothioate oligonucleotides may be

synthesized by the method of Stein et al. (1988, Nucl. Acids Res. 16:3209), methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin et al., 1988, Proc. Natl. Acad. Sci. U.S.A. 85:7448-7451), etc.

While antisense nucleotides complementary to the target gene coding region sequence could be used, those complementary to the transcribed untranslated region are most preferred.

Specific antisense oligonucleotides for the rchd534 gene (both short and long spliceoforms) and fchd540 gene are described in the example in Section 17, below.

The antisense molecules should be delivered to cells which express the target gene *in vivo*, e.g., endothelial cells. A number of methods have been developed for delivering antisense DNA or RNA to cells; e.g., antisense molecules can be injected directly into the tissue site, or modified antisense molecules, designed to target the desired cells (e.g., antisense linked to peptides or antibodies that specifically bind receptors or antigens expressed on the target cell surface) can be administered systemically.

However, it is often difficult to achieve intracellular concentrations of the antisense sufficient to suppress translation of endogenous mRNAs. Therefore a preferred approach utilizes a recombinant DNA construct in which the antisense oligonucleotide is placed under the control of a strong pol III or pol II promoter. The use of such a construct to transfect target cells in the patient will result in the transcription of sufficient amounts of single stranded RNAs that will form complementary base pairs with the endogenous target gene transcripts and thereby prevent translation of the target gene mRNA. For example, a vector can be introduced in vivo such that it is taken up by a cell and directs the transcription of an antisense RNA. Such a vector can remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the desired antisense RNA. Such vectors can be constructed by

recombinant DNA technology methods standard in the art. Vectors can be plasmid, viral, or others known in the art, used for replication and expression in mammalian cells. Expression of the sequence encoding the antisense RNA can be by any promoter known in the art to act in mammalian, preferably human cells. Such promoters can be inducible or constitutive. Such promoters include but are not limited to: the SV40 early promoter region (Bernoist and Chambon, 1981, Nature 290:304-310), the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto et al., 1980, Cell 22:787-797), the herpes thymidine kinase promoter (Wagner et al., 1981, Proc. Natl. Acad. Sci. U.S.A. 78:1441-1445), the regulatory sequences of the metallothionein gene (Brinster et al., 1982, Nature 296:39-42), etc. Any type of plasmid, cosmid, YAC or viral vector can be used to prepare the recombinant DNA construct which can be introduced directly into the tissue site; e.g., atherosclerotic vascular tissue. Alternatively, viral vectors can be used which selectively infect the desired tissue, in which case administration may be accomplished by another route (e.g., systemically).

Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. The mechanism of ribozyme action involves sequence specific hybridization of the ribozyme molecule to complementary target RNA, followed by an endonucleolytic cleavage. Ribozyme molecules designed to catalytically cleave target gene mRNA transcripts can also be used to prevent translation of target gene mRNA and expression of target gene. (See, e.g., PCT International Publication WO90/11364, published October 4, 1990; Sarver et al., 1990, Science 247:1222-1225). While ribozymes that cleave mRNA at site specific recognition sequences can be used to destroy target gene mRNAs, the use of hammerhead ribozymes is preferred. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. The sole requirement is that the target mRNA have the following

sequence of two bases: 5'-UG-3'. The construction and production of hammerhead ribozymes is well known in the art and is described more fully in Haseloff and Gerlach, 1988, Nature, 334:585-591. For example, there are hundreds of potential hammerhead ribozyme cleavage sites within the nucleotide sequence of rchd534 (both short and long spliceoforms) and fchd540 cDNA. Preferably the ribozyme is engineered so that the cleavage recognition site is located near the 5' end of the target mRNA; i.e., to increase efficiency and minimize the intracellular accumulation of non-functional mRNA transcripts.

Specific hammerhead ribozymes molecules for the rchd534 (both short and long spliceoforms) and fchd540 genes are described in the Example in Section 13, below.

The ribozymes of the present invention also include RNA endoribonucleases (hereinafter "Cech-type ribozymes") such as the one which occurs naturally in Tetrahymena Thermophila (known as the IVS, or L-19 IVS RNA) and which has been extensively described by Thomas Cech and collaborators (Zaug, et al., 1984, Science, 224:574-578; Zaug and Cech, 1986, Science, 231:470-475; Zaug, et al., 1986, Nature, 324:429-433; published International patent application No. WO 88/04300 by University Patents Inc.; Been and Cech, 1986, Cell, 47:207-216). The Cech-type ribozymes have an eight base pair active site which hybridizes to a target RNA sequence whereafter cleavage of the target RNA takes place. The invention encompasses those Cech-type ribozymes which target eight base-pair active site sequences that are present in target gene.

As in the antisense approach, the ribozymes can be composed of modified oligonucleotides (e.g. for improved stability, targeting, etc.) and should be delivered to cells which express the target gene *in vivo*, e.g., endothelial cells. A preferred method of delivery involves using a DNA construct "encoding" the ribozyme under the control of a strong constitutive pol III or pol II promoter, so that transfected cells will produce sufficient quantities of the

ribozyme to destroy endogenous target gene messages and inhibit translation. Because ribozymes, unlike antisense molecules, are catalytic, a lower intracellular concentration is required for efficiency.

Nucleic acid molecules to be used in triple helix formation for the inhibition of transcription should be single stranded and composed of deoxyribonucleotides. The base composition of these oligonucleotides must be designed to promote triple helix formation via Hoogsteen base pairing rules, which generally require sizeable stretches of either purines or pyrimidines to be present on one strand of a duplex. Nucleotide sequences may be pyrimidine-based, which will result in TAT and CGC^+ triplets across the three associated strands of the resulting triple helix. The pyrimidine-rich molecules provide base complementarity to a purine-rich region of a single strand of the duplex in a parallel orientation to that strand. In addition, nucleic acid molecules may be chosen that are purine-rich, for example, containing a stretch of G residues. These molecules will form a triple helix with a DNA duplex that is rich in GC paris, in which the majority of the purine residues are located on a single strand of the targeted duplex, resulting in GGC triplets across the three strands in the triplex.

Alternatively, the potential sequences that can be targeted for triple helix formation may be increased by creating a so called "switchback" nucleic acid molecule. Switchback molecules are synthesized in an alternating 5′-3′, 3′-5′ manner, such that they base pair with first one strand of a duplex and then the other, eliminating the necessity for a sizeable stretch of either purines or pyrimidines to be present on one strand of a duplex.

It is possible that the antisense, ribozyme, and/or triple helix molecules described herein may reduce or inhibit the transcription (triple helix) and/or translation (antisense, ribozyme) of mRNA produced by both normal and mutant target gene alleles. In order to ensure that substantially normal levels of target gene activity are

maintained, nucleic acid molecules that encode and express target gene polypeptides exhibiting normal activity may be introduced into cells via gene therapy methods such as those described, below, in Section 5.7. that do not contain sequences susceptible to whatever antisense, ribozyme, or triple helix treatments are being utilized. Alternatively, it may be preferable to coadminister normal target gene protein into the cell or tissue in order to maintain the requisite level of cellular or tissue target gene activity.

Endogenous target gene expression can also be reduced by inactivating or "knocking out" the target gene or its promoter using targeted homologous recombination. (E.g., see Smithies et al., 1985, Nature 317:230-234; Thomas & Capecchi, 1987, Cell 51:503-512; Thompson et al., 1989 Cell 5:313-321; each of which is incorporated by reference herein in its entirety). For example, a mutant, non-functional target (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous target gene (either the coding regions or regulatory regions of the target gene) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that express target *in vivo*. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the target gene. Such approaches can be adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site in vivo using appropriate viral vectors, e.g., vectors for delivery vascular tissue.

Alternatively, endogenous target gene expression can be reduced by targeting deoxyribonucleotide sequences complementary to the regulatory region of the target gene (i.e., the target promoter and/or enhancers) to form triple helical structures that prevent transcription of the target gene in target cells in the body. (See generally, Helene, C. 1991, Anticancer Drug Des., 6(6):569-84; Helene, C., et al., 1992, Ann, N.Y. Accad. Sci., 660:27-36; and Maher, L.J., 1992, Bioassays 14(12):807-15).

In yet another embodiment of the invention, the activity of a target can be reduced using a "dominant negative" approach to effectuate reduction in cardiovascular disease symptoms. For example, if two gene products interact, such as the rchd534 and fchd540 proteins, or the rchd534-long and fchd540 proteins, then the presence of a mutant version of one or both of these proteins in the cell can reduce the overall pool of complexes consisting of entirely wild-type proteins. In this manner, the overall level of activity resulting from the rchd534/fchd540 or rchd534-long/fchd540 protein interaction can be reduced.

5.6.1.2. ANTIBODIES FOR TARGET GENE PRODUCTS

Antibodies that are both specific for target gene protein and interfere with its activity may be used to inhibit target gene function. Such antibodies may be generated using standard techniques described in Section 5.4.3., *supra*, against the proteins themselves or against peptides corresponding to portions of the proteins. Such antibodies include but are not limited to polyclonal, monoclonal, Fab fragments, single chain antibodies, chimeric antibodies, etc.

In instances where the target gene protein is intracellular and whole antibodies are used, internalizing antibodies may be preferred. However, lipofectin liposomes may be used to deliver the antibody or a fragment of the Fab region which binds to the target gene epitope into cells. Where fragments of the antibody are used, the smallest inhibitory fragment which binds to the target protein's binding domain is preferred. For example, peptides having an amino acid sequence corresponding to the domain of the variable region of the antibody that binds to the target gene protein may be used. Such peptides may be synthesized chemically or produced via recombinant DNA technology using methods well known in the art (e.g., see Creighton, 1983, *supra*; and Sambrook et al., 1989, *supra*). Alternatively, single chain neutralizing antibodies which bind to

intracellular target gene epitopes may also be administered. Such single chain antibodies may be administered, for example, by expressing nucleotide sequences encoding single-chain antibodies within the target cell population by utilizing, for example, techniques such as those described in Marasco et al. (Marasco, W. et al., 1993, Proc. Natl. Acad. Sci. USA 90:7889-7893).

In some instances, the target gene protein is extracellular, or is a transmembrane protein, such as the rchd523 gene product. Antibodies that are specific for one or more extracellular domains of the rchd523 gene product, for example, and that interfere with its activity, are particularly useful in treating cardiovascular disease. Such antibodies are especially efficient because they can access the target domains directly from the bloodstream. Any of the administration techniques described, below in Section 5.7 which are appropriate for peptide administration may be utilized to effectively administer inhibitory target gene antibodies to their site of action.

5.6.2. METHODS FOR RESTORING OR ENHANCING TARGET GENE ACTIVITY

Target genes that cause cardiovascular disease may be underexpressed within cardiovascular disease situations. As summarized in Table 1, above, and detailed in the example in Sections 7, below, several genes are now known to be down-regulated in monocytes under disease conditions. Specifically, bcl-2 and glutathione peroxidase gene expression is down-regulated in the monocytes of patients exposed to a high lipid diet, e.g. cholesterol or fat, that leads to high serum LDL levels. Alternatively, the activity of target gene products may be decreased, leading to the development of cardiovascular disease symptoms. Such down-regulation of target gene expression or decrease of target gene product activity might have a causative or exacerbating effect on the disease state.

In some cases, target genes that are up-regulated in the disease state might be exerting a protective effect. As summarized in Table 1, above, and detailed in the examples in Sections 8 and 9, below, a number of genes are now known to be up-regulated in endothelial cells under disease conditions. Specifically, rchd005, rchd024, rchd032, and rchd036 are all up-regulated in endothelial cells treated with IL-1. Furthermore, rchd502, rchd523, rchd528, rchd534, COX II, and MnSOD are all up-regulated in endothelial cells subjected to shear stress. A variety of techniques may be utilized to increase the expression, synthesis, or activity of such target genes and/or proteins, for those genes that exert a protective effect in response to disease conditions.

Described in this Section are methods whereby the level of target gene activity may be increased to levels wherein cardiovascular disease symptoms are ameliorated. The level of gene activity may be increased, for example, by either increasing the level of target gene product present or by increasing the level of active target gene product which is present.

For example, a target gene protein, at a level sufficient to ameliorate cardiovascular disease symptoms may be administered to a patient exhibiting such symptoms. Any of the techniques discussed, below, in Section 5.7, may be utilized for such administration. One of skill in the art will readily know how to determine the concentration of effective, non-toxic doses of the normal target gene protein, utilizing techniques such as those described, below, in Section 5.7.1.

Additionally, RNA sequences encoding target gene protein may be directly administered to a patient exhibiting cardiovascular disease symptoms, at a concentration sufficient to produce a level of target gene protein such that cardiovascular disease symptoms are ameliorated. Any of the techniques discussed, below, in Section 5.7, which achieve intracellular administration of compounds, such as, for example, liposome administration, may be utilized for the

administration of such RNA molecules. The RNA molecules may be produced, for example, by recombinant techniques such as those described, above, in Section 5.4.2.

Further, patients may be treated by gene replacement therapy. One or more copies of a normal target gene, or a portion of the gene that directs the production of a normal target gene protein with target gene function, may be inserted into cells using vectors which include, but are not limited to adenovirus, adeno-associated virus, and retrovirus vectors, in addition to other particles that introduce DNA into cells, such as liposomes. Additionally, techniques such as those described above may be utilized for the introduction of normal target gene sequences into human cells.

Cells, preferably, autologous cells, containing normal target gene expressing gene sequences may then be introduced or reintroduced into the patient at positions which allow for the amelioration of cardiovascular disease symptoms. Such cell replacement techniques may be preferred, for example, when the target gene product is a secreted, extracellular gene product.

5.7. PHARMACEUTICAL PREPARATIONS AND METHODS OF ADMINISTRATION

The identified compounds that inhibit target gene expression, synthesis and/or activity can be administered to a patient at therapeutically effective doses to treat or ameliorate cardiovascular disease. A therapeutically effective dose refers to that amount of the compound sufficient to result in amelioration of symptoms of cardiovascular disease.

5.7.1. EFFECTIVE DOSE

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD_{50} (the dose lethal to 50% of the population) and the

ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD_{50}/ED_{50}. Compounds which exhibit large therapeutic indices are preferred. While compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC_{50} (i.e., the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

5.7.2. FORMULATIONS AND USE

Pharmaceutical compositions for use in accordance with the present invention may be formulated in conventional manner using one or more physiologically acceptable carriers or excipients.

Thus, the compounds and their physiologically acceptable salts and solvates may be formulated for administration by inhalation or insufflation (either through the mouth or the nose) or oral, buccal, parenteral or rectal administration.

For oral administration, the pharmaceutical compositions may take the form of, for example, tablets or capsules prepared by conventional means with pharmaceutically acceptable excipients such as binding agents (e.g., pregelatinised maize starch, polyvinylpyrrolidone or hydroxypropyl methylcellulose); fillers (e.g., lactose, microcrystalline cellulose or calcium hydrogen phosphate); lubricants (e.g., magnesium stearate, talc or silica); disintegrants (e.g., potato starch or sodium starch glycolate); or wetting agents (e.g., sodium lauryl sulphate). The tablets may be coated by methods well known in the art. Liquid preparations for oral administration may take the form of, for example, solutions, syrups or suspensions, or they may be presented as a dry product for constitution with water or other suitable vehicle before use. Such liquid preparations may be prepared by conventional means with pharmaceutically acceptable additives such as suspending agents (e.g., sorbitol syrup, cellulose derivatives or hydrogenated edible fats); emulsifying agents (e.g., lecithin or acacia); non-aqueous vehicles (e.g., almond oil, oily esters, ethyl alcohol or fractionated vegetable oils); and preservatives (e.g., methyl or propyl-p-hydroxybenzoates or sorbic acid). The preparations may also contain buffer salts, flavoring, coloring and sweetening agents as appropriate.

Preparations for oral administration may be suitably formulated to give controlled release of the active compound.

For buccal administration the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon

dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of e.g. gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch.

The compounds may be formulated for parenteral administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampoules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, e.g., containing conventional suppository bases such as cocoa butter or other glycerides.

In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may for example comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration.

5.8. DIAGNOSIS OF CARDIOVASCULAR DISEASE ABNORMALITIES

A variety of methods may be employed, utilizing reagents such as fingerprint gene nucleotide sequences described in Section 5.4.1, and antibodies directed against differentially expressed and pathway gene peptides, as described, above, in Sections 5.4.2. (peptides) and 5.4.3. (antibodies). Specifically, such reagents may be used, for example, for the detection of the presence of target gene mutations, or the detection of either over or under expression of target gene mRNA.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one specific fingerprint gene nucleic acid or anti-fingerprint gene antibody reagent described herein, which may be conveniently used, e.g., in clinical settings, to diagnose patients exhibiting cardiovascular disease symptoms or at risk for developing cardiovascular disease.

Any cell type or tissue, preferably monocytes, endothelial cells, or smooth muscle cells, in which the fingerprint gene is expressed may be utilized in the diagnostics described below.

5.8.1. DETECTION OF FINGERPRINT GENE NUCLEIC ACIDS

DNA or RNA from the cell type or tissue to be analyzed may easily be isolated using procedures which are well known to those in the art. Diagnostic procedures may also be performed "in situ" directly upon tissue sections (fixed and/or frozen) of patient tissue obtained from biopsies or resections, such that no nucleic acid purification is necessary. Nucleic acid reagents such as those described in Section 5.1. may be used as probes and/or primers for such in situ procedures (see, for example, Nuovo, G.J., 1992, PCR in situ hybridization: protocols and applications, Raven Press, NY).

Fingerprint gene nucleotide sequences, either RNA or DNA, may, for example, be used in hybridization or

amplification assays of biological samples to detect cardiovascular disease-related gene structures and expression. Such assays may include, but are not limited to, Southern or Northern analyses, single stranded conformational polymorphism analyses, in situ hybridization assays, and polymerase chain reaction analyses. Such analyses may reveal both quantitative aspects of the expression pattern of the fingerprint gene, and qualitative aspects of the fingerprint gene expression and/or gene composition. That is, such aspects may include, for example, point mutations, insertions, deletions, chromosomal rearrangements, and/or activation or inactivation of gene expression.

Such an in situ hybridization analysis is described in the example in Section 14, below. Specifically, high levels of expression of the rchd502 and rchd528 genes were detected specifically within the endothelial cells of diseased tissue removed from a human cardiovascular disease patient, and not in any other cell type present in the tissue, including smooth muscle cells and macrophages. These results clearly demonstrate how the target genes described herein provide for novel diagnoses of cardiovascular disease. Furthermore, because these diagnoses are correlated with specific target genes, they allow for more specifically directed methods of treatment of cardiovascular disease.

Preferred diagnostic methods for the detection of fingerprint gene-specific nucleic acid molecules may involve for example, contacting and incubating nucleic acids, derived from the cell type or tissue being analyzed, with one or more labeled nucleic acid reagents as are described in Section 5.1, under conditions favorable for the specific annealing of these reagents to their complementary sequences within the nucleic acid molecule of interest. Preferably, the lengths of these nucleic acid reagents are at least 9 to 30 nucleotides. After incubation, all non-annealed nucleic acids are removed from the nucleic acid:fingerprint molecule hybrid. The presence of nucleic acids from the fingerprint tissue which have hybridized, if any such molecules exist, is

then detected. Using such a detection scheme, the nucleic acid from the tissue or cell type of interest may be immobilized, for example, to a solid support such as a membrane, or a plastic surface such as that on a microtitre plate or polystyrene beads. In this case, after incubation, non-annealed, labeled fingerprint nucleic acid reagents of the type described in Section 5.1. are easily removed. Detection of the remaining, annealed, labeled nucleic acid reagents is accomplished using standard techniques well-known to those in the art.

Alternative diagnostic methods for the detection of fingerprint gene specific nucleic acid molecules may involve their amplification, e.g., by PCR (the experimental embodiment set forth in Mullis, K.B., 1987, U.S. Patent No. 4,683,202), ligase chain reaction (Barany, F., 1991, Proc. Natl. Acad. Sci. USA 88:189-193), self sustained sequence replication (Guatelli, J.C. et al., 1990, Proc. Natl. Acad. Sci. USA 87:1874-1878), transcriptional amplification system (Kwoh, D.Y et al., 1989, Proc. Natl. Acad. Sci. USA 86:1173-1177), Q-Beta Replicase (Lizardi, P.M. et al., 1988, Bio/Technology 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In one embodiment of such a detection scheme, a cDNA molecule is obtained from an RNA molecule of interest (e.g., by reverse transcription of the RNA molecule into cDNA). Cell types or tissues from which such RNA may be isolated include any tissue in which wild type fingerprint gene is known to be expressed, including, but not limited, to monocytes, endothelium, and/or smooth muscle. A fingerprint sequence within the cDNA is then used as the template for a nucleic acid amplification reaction, such as a PCR amplification reaction, or the like. The nucleic acid reagents used as synthesis initiation reagents (e.g.,

primers) in the reverse transcription and nucleic acid amplification steps of this method are chosen from among the fingerprint gene nucleic acid reagents described in Section 5.1. The preferred lengths of such nucleic acid reagents are at least 15-30 nucleotides. For detection of the amplified product, the nucleic acid amplification may be performed using radioactively or non-radioactively labeled nucleotides. Alternatively, enough amplified product may be made such that the product may be visualized by standard ethidium bromide staining or by utilizing any other suitable nucleic acid staining method.

In addition to methods which focus primarily on the detection of one nucleic acid sequence, fingerprint profiles, as discussed in Section 5.5.4, may also be assessed in such detection schemes. Fingerprint profiles may be generated, for example, by utilizing a differential display procedure, as discussed, above, in Section 5.1.2, Northern analysis and/or RT-PCR. Any of the gene sequences described, above, in Section 5.4.1. may be used as probes and/or PCR primers for the generation and corroboration of such fingerprint profiles.

5.8.2. DETECTION OF FINGERPRINT GENE PEPTIDES

Antibodies directed against wild type or mutant fingerprint gene peptides, which are discussed, above, in Section 5.4.3, may also be used as cardiovascular disease diagnostics and prognostics, as described, for example, herein. Such diagnostic methods, may be used to detect abnormalities in the level of fingerprint gene protein expression, or abnormalities in the structure and/or tissue, cellular, or subcellular location of fingerprint gene protein. Structural differences may include, for example, differences in the size, electronegativity, or antigenicity of the mutant fingerprint gene protein relative to the normal fingerprint gene protein.

Protein from the tissue or cell type to be analyzed may easily be detected or isolated using techniques which are

well known to those of skill in the art, including but not limited to western blot analysis. For a detailed explanation of methods for carrying out western blot analysis, see Sambrook et al, 1989, supra, at Chapter 18. The protein detection and isolation methods employed herein may also be such as those described in Harlow and Lane, for example, (Harlow, E. and Lane, D., 1988, "Antibodies: A Laboratory Manual", Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York), which is incorporated herein by reference in its entirety.

Preferred diagnostic methods for the detection of wild type or mutant fingerprint gene peptide molecules may involve, for example, immunoassays wherein fingerprint gene peptides are detected by their interaction with an anti-fingerprint gene specific peptide antibody.

For example, antibodies, or fragments of antibodies, such as those described, above, in Section 5.4.3, useful in the present invention may be used to quantitatively or qualitatively detect the presence of wild type or mutant fingerprint gene peptides. This can be accomplished, for example, by immunofluorescence techniques employing a fluorescently labeled antibody (see below) coupled with light microscopic, flow cytometric, or fluorimetric detection. Such techniques are especially preferred if the fingerprint gene peptides are expressed on the cell surface.

The antibodies (or fragments thereof) useful in the present invention may, additionally, be employed histologically, as in immunofluorescence or immunoelectron microscopy, for in situ detection of fingerprint gene peptides. In situ detection may be accomplished by removing a histological specimen from a patient, and applying thereto a labeled antibody of the present invention. The antibody (or fragment) is preferably applied by overlaying the labeled antibody (or fragment) onto a biological sample. Through the use of such a procedure, it is possible to determine not only the presence of the fingerprint gene peptides, but also their distribution in the examined tissue. Using the present

invention, those of ordinary skill will readily perceive that any of a wide variety of histological methods (such as staining procedures) can be modified in order to achieve such in situ detection.

Immunoassays for wild type or mutant fingerprint gene peptides typically comprise incubating a biological sample, such as a biological fluid, a tissue extract, freshly harvested cells, or cells which have been incubated in tissue culture, in the presence of a detectably labeled antibody capable of identifying fingerprint gene peptides, and detecting the bound antibody by any of a number of techniques well known in the art.

The biological sample may be brought in contact with and immobilized onto a solid phase support or carrier such as nitrocellulose, or other solid support which is capable of immobilizing cells, cell particles or soluble proteins. The support may then be washed with suitable buffers followed by treatment with the detectably labeled fingerprint gene specific antibody. The solid phase support may then be washed with the buffer a second time to remove unbound antibody. The amount of bound label on solid support may then be detected by conventional means.

By "solid phase support or carrier" is intended any support capable of binding an antigen or an antibody. Well-known supports or carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, gabbros, and magnetite. The nature of the carrier can be either soluble to some extent or insoluble for the purposes of the present invention. The support material may have virtually any possible structural configuration so long as the coupled molecule is capable of binding to an antigen or antibody. Thus, the support configuration may be spherical, as in a bead, or cylindrical, as in the inside surface of a test tube, or the external surface of a rod. Alternatively, the surface may be flat such as a sheet, test strip, etc. Preferred supports include polystyrene beads. Those skilled

in the art will know many other suitable carriers for binding antibody or antigen, or will be able to ascertain the same by use of routine experimentation.

The binding activity of a given lot of anti-wild type or mutant fingerprint gene peptide antibody may be determined according to well known methods. Those skilled in the art will be able to determine operative and optimal assay conditions for each determination by employing routine experimentation.

One of the ways in which the fingerprint gene peptide-specific antibody can be detectably labeled is by linking the same to an enzyme and use in an enzyme immunoassay (EIA) (Voller, "The Enzyme Linked Immunosorbent Assay (ELISA)", *Diagnostic Horizons* 2:1-7, 1978, Microbiological Associates Quarterly Publication, Walkersville, MD; Voller, et al., J. Clin. Pathol. 31:507-520 (1978); Butler, Meth. Enzymol. 73:482-523 (1981); Maggio, (ed.) *Enzyme Immunoassay*, CRC Press, Boca Raton, FL, 1980; Ishikawa, et al., (eds.) *Enzyme Immunoassay*, Kgaku Shoin, Tokyo, 1981). The enzyme which is bound to the antibody will react with an appropriate substrate, preferably a chromogenic substrate, in such a manner as to produce a chemical moiety which can be detected, for example, by spectrophotometric, fluorimetric or by visual means. Enzymes which can be used to detectably label the antibody include, but are not limited to, malate dehydrogenase, staphylococcal nuclease, delta-5-steroid isomerase, yeast alcohol dehydrogenase, alpha-glycerophosphate, dehydrogenase, triose phosphate isomerase, horseradish peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, beta-galactosidase, ribonuclease, urease, catalase, glucose-6-phosphate dehydrogenase, glucoamylase and acetylcholinesterase. The detection can be accomplished by colorimetric methods which employ a chromogenic substrate for the enzyme. Detection may also be accomplished by visual comparison of the extent of enzymatic reaction of a substrate in comparison with similarly prepared standards.

Detection may also be accomplished using any of a variety of other immunoassays. For example, by radioactively labeling the antibodies or antibody fragments, it is possible to detect fingerprint gene wild type or mutant peptides through the use of a radioimmunoassay (RIA) (see, for example, Weintraub, B., *Principles of Radioimmunoassays*, Seventh Training Course on Radioligand Assay Techniques, The Endocrine Society, March, 1986, which is incorporated by reference herein). The radioactive isotope can be detected by such means as the use of a gamma counter or a scintillation counter or by autoradiography.

It is also possible to label the antibody with a fluorescent compound. When the fluorescently labeled antibody is exposed to light of the proper wave length, its presence can then be detected due to fluorescence. Among the most commonly used fluorescent labeling compounds are fluorescein isothiocyanate, rhodamine, phycoerythrin, phycocyanin, allophycocyanin, o-phthaldehyde and fluorescamine.

The antibody can also be detectably labeled using fluorescence emitting metals such as ^{152}Eu, or others of the lanthanide series. These metals can be attached to the antibody using such metal chelating groups as diethylenetriaminepentacetic acid (DTPA) or ethylenediaminetetraacetic acid (EDTA).

The antibody also can be detectably labeled by coupling it to a chemiluminescent compound. The presence of the chemiluminescent-tagged antibody is then determined by detecting the presence of luminescence that arises during the course of a chemical reaction. Examples of particularly useful chemiluminescent labeling compounds are luminol, isoluminol, theromatic acridinium ester, imidazole, acridinium salt and oxalate ester.

Likewise, a bioluminescent compound may be used to label the antibody of the present invention. Bioluminescence is a type of chemiluminescence found in biological systems in, which a catalytic protein increases the efficiency of the

chemiluminescent reaction. The presence of a bioluminescent protein is determined by detecting the presence of luminescence. Important bioluminescent compounds for purposes of labeling are luciferin, luciferase and aequorin.

5.8.3. IMAGING CARDIOVASCULAR DISEASE CONDITIONS

In some cases, differentially expressed gene products identified herein may be up-regulated under cardiovascular disease conditions and expressed on the surface of the affected tissue. Such target gene products allow for the non-invasive imaging of damaged or diseased cardiovascular tissue for the purposed of diagnosis and directing of treatment of the disease. For example, such differentially expressed gene products may include but are not limited to atherosclerosis specific adhesion molecules responsible for atherogenesis, or monocyte scavenger receptors that are up-regulated in response to oxidized LDL, which are discussed in Section 2, above. Alternatively, other such surface proteins may be specifically up-regulated in tissues suffering from ischemia/reperfusion or other tissues with atherosclerotic or restenotic lesions.

As described in the example in Section 9, below, the rchd523 gene is a gene that is up-regulated in endothelial cells under shear stress. Furthermore, the rchd523 gene encodes a novel G protein-coupled receptor, containing an extracellular amino terminal domain, in addition to seven transmembrane domains. The rchd523 gene product, therefore, provides an excellent tool for imaging cardiovascular disease conditions. This technique can be applied similarly to other transmembrane gene products, such as the rchd502 and rchd528 gene products. An example illustrating the use of this method in accordance with the invention is provided in Section 11, below.

Monoclonal antibodies, as described in Section 5.6.1.2, above, which specifically bind to such surface proteins, such as the rchd523 gene product, may be used for the diagnosis of cardiovascular disease by in vivo tissue

imaging techniques. An antibody specific for a target gene product, or preferably an antigen binding fragment thereof, is conjugated to a label (e.g., a gamma emitting radioisotope) which generates a detectable signal and administered to a subject (human or animal) suspected of having cardiovascular disease. After sufficient time to allow the detectably-labeled antibody to localize at the diseased or damaged tissue site (or sites), the signal generated by the label is detected by a photoscanning device. The detected signal is then converted to an image of the tissue. This image makes it possible to localize the tissue in vivo. This data can then be used to develop an appropriate therapeutic strategy.

Antibody fragments, rather than whole antibody molecules, are generally preferred for use in tissue imaging. Antibody fragments accumulate at the tissue(s) more rapidly because they are distributed more readily than are entire antibody molecules. Thus an image can be obtained in less time than is possible using whole antibody. These fragments are also cleared more rapidly from tissues, resulting in a lower background signal. See, e.g., Haber et al., U.S. Patent No. 4,036,945; Goldenberg et al., U.S. Patent No. 4,331,647. The divalent antigen binding fragment $(Fab')_2$ and the monovalent Fab are especially preferred. Such fragments can be prepared by digestion of the whole immunoglobulin molecule with the enzymes pepsin or papain according to any of several well known protocols. The types of labels that are suitable for conjugation to a monoclonal antibody for diseased or damaged tissue localization include, but are not limited to radiolabels (i.e., radioisotopes), fluorescent labels and biotin labels.

Among the radioisotopes that can be used to label antibodies or antibody fragments, gamma-emitters, positron-emitters, X-ray-emitters and fluorescence-emitters are suitable for localization. Suitable radioisotopes for labeling antibodies include Iodine-131, Iodine-123, Iodine-125, Iodine-126, Iodine-133, Bromine-77, Indium-111, Indium-

113m, Gallium-67, Gallium-68, Ruthenium-95, Ruthenium-97, Ruthenium-103, Ruthenium-105, Mercury-107, Mercury-203, Rhenium-99m, Rhenium-105, Rhenium-101, Tellurium-121m, Tellurium-122m, Tellurium-125m, Thulium-165, Thulium-167, Thulium-168, Technetium-99m and Fluorine-18. The halogens can be used more or less interchangeably as labels since halogen-labeled antibodies and/or normal immunoglobulins would have substantially the same kinetics and distribution and similar metabolism.

The gamma-emitters Indium-111 and Technetium-99m are preferred because these radiometals are detectable with a gamma camera and have favorable half lives for imaging in vivo. Antibody can be labelled with Indium-111 or Technetium-99m via a conjugated metal chelator, such as DTPA (diethlenetriaminepentaacetic acid). See Krejcarek et al., 1977, Biochem. Biophys. Res. Comm. 77:581; Khaw et al., 1980, Science 209:295; Gansow et al., U.S. Patent No. 4,472,509; Hnatowich, U.S. Patent No. 4,479,930, the teachings of which are incorporated herein by reference.

Fluorescent compounds that are suitable for conjugation to a monoclonal antibody include fluorescein sodium, fluorescein isothiocyanate, and Texas Red sulfonyl chloride. See, DeBelder & Wik, 1975, Carbohydrate Research 44:254-257. Those skilled in the art will know, or will be able to ascertain with no more than routine experimentation, other fluorescent compounds that are suitable for labeling monoclonal antibodies.

6. EXAMPLE: IDENTIFICATION OF GENES DIFFERENTIALLY EXPRESSED IN RESPONSE TO PARADIGM A: IN VITRO FOAM CELL PARADIGM

According to the invention, differential display may be used to detect genes that are differentially expressed in monocytes that were treated so as to simulate the conditions under which foam cells develop during atherogenesis.

6.1. MATERIALS AND METHODS

6.1.1. CELL ISOLATION AND CULTURING

Blood (~200 ml) was drawn into chilled 20 ml vacutainer tubes to which 3 ml of citrate phosphate dextrose (Sigma) was added. Blood was then pooled into 50 ml tubes and spun in the Beckman GS-6R at 1250 RPM for 15 minutes at 4°C. The upper clear layer (~25 ml) was then removed with a pipette and discarded and replaced with the same volume of 4°C PBS. The blood was then mixed, and spun again at 2680 RPM for 15 minutes at 4°C. The upper layer was then removed and discarded, and the buffy coat at the interface was removed in ~5 ml and placed in a separate 50 ml tube, and the pipette was washed with 20 ml PBS. Cells were added to a T flask and stored at 4°C for 16 hours. A small aliquot of the cells were then removed and counted using a hemacytometer. The final red blood cell concentration in the buffy coat population was then adjusted to 1.5 X 10^9/ml with PBS, the cells were added to Leucoprep tubes (Becton Dickinson) after being allowed to come to room temperature, and spun at 2300 RPM for 25 minutes at 25°C. The upper clear layer was removed and discarded and the turbid layer over the gel was removed and pooled in 50 ml tubes. Samples were then diluted to 50 ml with PBS (25°C) and spun at 1000 RPM for 10 minutes. The supernatant was then removed, and the pellet was resuspended in 50 ml PBS. This procedure was repeated 3 more times. After the last spin, the cells were resuspended in a small volume of PBS and counted.

Tissue culture dishes were coated with bovine collagen before monocytes were plated out. 1/6 volume of 7X RPMI (JRH Biosciences) was added to Vitrogen 100 collagen (Celtrix) which was then diluted 1:10 with RPMI to a final concentration of 0.35 mg/ml. Collagen mixture was then added to plates (2.5 ml/100 mm dish) and placed at 37°C for at least one hour to allow for gel formation. After gel formation has taken place, the RPMI was removed and cells were added in RPMI/10% plasma derived serum (PDS). PDS was prepared by drawing blood into chilled evacuated tubes

containing 1/10th volume 3.8% sodium citrate. Blood was then transferred into new Sorvall tubes and spun at 14,000-16,000 RPM for 20 minutes at 4°C. Plasma layer was removed and pooled in new tubes to which 1/50th volume 1M $CaCl_2$ was added. Plasma was mixed and aliquoted into new Sorvall tubes and incubated at 37° for 2 hours to allow for fibrin clot formation. The clot was then disturbed with a pipette to allow it to contract and tubes were spun at 14,500 RPM for 20 minutes at 25°C. Supernatant was collected, pooled, and heat inactivated at 56°C prior to sterile filtration and freezing.

Purified human monocytes were cultured in 10% PDS/RPMI containing 5 units/ml of Genzyme recombinant human MCSF for 5 days before being treated with LDL, oxidized LDL, acetylated LDL (all LDL at 50 μg/ml), lysophosphatidylcholine (Sigma, 37.5 μM), or homocysteine (Sigma, 1mM). After incubation with these reagents for periods ranging from 2 hours up to 3 days, the media was withdrawn and the cells were dissolved in RNA lysis buffer and RNA was prepared as described, above, in Section 6.1.

Lipoproteins For oxidation, human LDL (Sigma) was first diluted to 1 mg/ml with PBS and then dialyzed against PBS at 4°C overnight. LDL was then diluted to 0.3 mg/ml with PBS. $CuSO_4 \cdot 5H_2O$ was then added to 5uM final concentration, and the solution was incubated in a T flask in a 37°C incubator for 24 hr. LDL solution was then dialyzed at 4°C against 0.15M NaCl/0.3mM EDTA for 2 days with several changes, before being removed and concentrated using an Amicon spin column by spinning for 1 hr. 4000 RPM at 4°C.

For acetylation, 1 ml of 5 mg/ml LDL was added to 1 ml of a saturated solution of NaOAc in a 15 ml tube on ice on a shaker at 4°C. 8 μl of acetic anhydride was added 2 μl at a time over 1 hr. LDL was then dialyzed for 48 hr. against 0.15M NaCl/0.3 mM EDTA at 4°C for 48 hr. with several changes. Final concentrations of derivatized LDL's were determined by comparing to a dilution curve of native LDL analyzed at OD_{280}, with 0.15M NaCl/0.3mM EDTA used as diluent in all cases.

6.1.2. ANALYSIS OF PARADIGM MATERIAL

Differential Display:

Removal of DNA: The RNA pellet was resuspended in H_2O and quantified by spectrophotometry at OD_{260}. Approximately half of the sample was then treated with DNAse I to remove contaminating chromosomal DNA. RNA was amplified by PCR using the following procedure. 50 ul RNA sample (10-20 μg), 5.7 μl 10x PCR buffer (Perkin-Elmer/Cetus), 1 μl RNAse inhibitor (40 units/μl) (Boehringer Mannheim, Germany) were mixed together, vortexed, and briefly spun. 2 μl DNAse I (10 units/μl) (Boehringer Mannheim) was added to the reaction which was incubated for 30 min. at 37°C. The total volume was brought to 200 μl with DEPC H_2O, extracted once with phenol/chloroform, once with chloroform, and precipitated by adding 20 μl 3M NaOAc, pH 4.8, (DEPC-treated), 500 μl absolute ETOH and incubating for 1 hour on dry ice or -20°C overnight. The precipitated sample was centrifuged for 15 min., and the pellet was washed with 70% ETOH. The sample was re-centrifuged, the remaining liquid was aspirated, and the pellet was resuspended in 100 μl H_2O. The concentration of RNA was measured by reading the OD_{260}.

First strand cDNA synthesis: For each RNA sample duplicate reactions were carried out in parallel. 400 ng RNA plus DEPC H_2O in a total volume of 10 μl were added to 4 μl $T_{11}XX$ reverse primer (10 μM) (Operon). The mixture was incubated at 70°C for 5 min. to denature the RNA and then placed at r.t. 26 μl of reaction mix containing the following components was added to each denatured RNA/primer sample: 8 μl 5x First Strand Buffer (Gibco/BRL, Gaithersburg, MD), 4 μl O.1M DTT (Gibco/BRL), 2 μl RNAse inhibitor (40 units/μl) (Boehringer Mannheim), 4 μl 200 μM dNTP mix, 6 μl H_2O, 2 μl Superscript reverse transcriptase (200 units/μl) (Gibco/BRL). The reactions were mixed gently and incubated for 30 min. at 42°C. 60 μl of H_2O (final volume = 100 μl) were then added and the samples were denatured for 5 min. at 85°C and stored at -20°C.

PCR reactions: 13 μl of reaction mix was added to each tube of a 96 well plate on ice. The reaction mix contained 6.4 μl H_2O, 2 μl 10x PCR Buffer (Perkin-Elmer), 2 μl 20 μM dNTP's, 0.4 μl ^{35}S dATP (12.5 μCi/μl; 50 μCi total) (Dupont/NEN), 2 μl forward primer (10 μM) (Operon), and 0.2 μl AmpliTaq Polymerase (5 units/μl) (Perkin-Elmer). Next, 2 μl of reverse primer ($T_{11}XX$, 10 μM) were added to the side of each tube followed by 5 μl of cDNA also to the sides of the tubes, which were still on ice. Tubes were capped and mixed, and brought up to 1000 RPM in a centrifuge then returned immediately to ice. The PCR machine (Perkin-Elmer 9600) was programmed for differential display as follows:

	94°C	2 min.
	*94°C	15 sec.
	*40°C	2 min.
* = X40	*ramp 72°C	1 min.
	*72°C	30 sec.
	72°C	5 min.
	4°C	hold

When the PCR machine reached 94°C, the plate was removed from ice and placed directly into the Perkin-Elmer 9600 PCR machine . Following PCR, 15 μl of loading dye, containing 80% formamide, 10 mM EDTA, 1 mg/ml xylene cyanol, 1 mg/ml bromphenol blue were added. The loading dye and reaction were mixed, incubated at 85°C for 5 min., cooled on ice, centrifuged, and placed on ice. Approximately 4 μl from each tube were loaded onto a prerun (60V) 6% acrylamide gel. The gel was run at approximately 80V until top dye front was about 1 inch from bottom. The gel was transferred to 3MM paper (Whatman Paper, England) and dried under vacuum. Bands were visualized by autoradiography.

Band isolation and amplification: Differentially expressed bands were excised from the dried gel with a razor blade and placed into a microfuge tube with 100 μl H_2O and heated at 100°C for 5 min., vortexed, heated again to 100°C for 5 min.,

and vortex again. After cooling, 100 μl H_2O, 20 μl 3M NaOAc, 1 μl glycogen (20 mg/ml), and 500 μl ethanol were added and chilled. After centrifugation, the pellet was washed and resuspended in 10 μl H_2O.

The isolated differentially expressed bands were then amplified by PCR using the following reaction conditions:

58	μl	H_2O
10	μl	10x PCR Buffer
10	μl	200 μm dNTP's
10	μl	10 μM reverse primer
10	μl	10 μM forward primer
1.5	μl	amplified band
0.5	μl	AmpliTaq polymerase (5 units/μl) (Perkin Elmer)

PCR was performed using the program described in this Section, above, for differential display. After PCR, glycerol loading dyes were added and samples were loaded onto a 2% preparative TAE/Biogel (Bio101, La Jolla, CA) agarose gel and eluted. Bands were then excised from the gel with a razor blade and vortexed for 15 min. at r.t., and purified using the Mermaid kit from Bio101 by adding 3 volumes of Mermaid high salt binding solution and 8 μl of resuspended glassfog in a microfuge tube. Glassfog was then pelleted, washed 3 times with ethanol wash solution, and then DNA was eluted twice in 10 μl at 50°C.

Subcloning: The TA cloning kit (Invitrogen, San Diego, CA) was used to subclone the amplified bands. The ligation reaction typically consisted of 4 μl sterile H_2O, 1 μl ligation buffer, 2 μl TA cloning vector, 2 μl PCR product, and 1 μl T4 DNA ligase. The volume of PCR product can vary, but the total volume of PCR product plus H_2O was always 6 μl. Ligations (including vector alone) were incubated overnight at 12°C before bacterial transformation. TA cloning kit competent bacteria (INVαF': *endA1*, *recA1*, *hsdR17*(r-k, m+k), *supE44*, λ-, *thi-1*, *gyrA*, *relA1*, φ80lacZαΔM15Δ(*lacZYA-argF*), deoR+, F') were thawed on ice and 2 μl of 0.5 M β-

mercaptoethanol were added to each tube. 2 μl from each ligation were added to each tube of competent cells (50 μl), mixed without vortexing, and incubated on ice for 30 min. Tubes were then placed in 42°C bath for exactly 30 sec., before being returned to ice for 2 min. 450 μl of SOC media (Sambrook et al., 1989, *supra*) were then added to each tube which were then shaken at 37°C for 1 hr. Bacteria were then pelleted, resuspended in ~200 μl SOC and plated on Luria broth agar plates containing X-gal and 60 μg/ml ampicillin and incubated overnight at 37°C. White colonies were then picked and screened for inserts using PCR.

A master mix containing 2 μl 10x PCR buffer, 1.6 μl 2.5 mM dNTP's, 0.1 μl 25 mM $MgCl_2$, 0.2 μl M13 reverse primer (100 ng/μl), 0.2 μl M13 forward primer (100 ng/μl), 0.1 μl AmpliTaq (Perkin-Elmer), and 15.8 μl H_20 was made. 40 μl of the master mix were aliquoted into tubes of a 96 well plate, and whole bacteria were added with a pipette tip prior to PCR. The PCR machine (Perkin-Elmer 9600) was programmed for insert screening as follows:

	94°C	2 min.
	*94°C	15 sec.
	*47°C	2 min.
* = X35	*ramp 72°C	30 sec.
	*72°C	30 sec.
	72°C	10 min.
	4°C	hold

Reaction products were eluted on a 2% agarose gel and compared to vector control. Colonies with vectors containing inserts were purified by streaking onto LB/Amp plates. Vectors were isolated from such strains and subjected to sequence analysis, using an Applied Biosystems Automated Sequencer (Applied Biosystems, Inc. Seattle, WA).

Northern analysis: Northern analysis was performed to confirm the differential expression of the genes corresponding to the amplified bands. The probes used to detect mRNA were synthesized as follows: typically 2 μl

amplified band (~30 ng), 7 μl H_2O, and 2 μl 10x Hexanucleotide mix (Boehringer-Mannheim) were mixed and heated to 95°C for 5 min., and then allowed to cool on ice. The volume of the amplified band can vary, but the total volume of the band plus H_2O was always 9 μl. 3 μl dATP/dGTP/dTTP mix (1:1:1 of 0.5 mM each), 5 μl α^{32}P dCTP 3000 Ci/mM (50 μCi total) (Amersham, Arlington Heights, IL), and 1 μl Klenow (2 units) (Boehringer-Mannheim) were mixed and incubated at 37°C. After 1 hr., 30 μl TE were added and the reaction was loaded onto a Biospin-6™ column (Biorad, Hercules, CA), and centrifuged. A 1 μl aliquot of eluate was used to measure incorporation in a scintillation counter with scintillant to ensure that 10^6cpm/μl of incorporation was achieved.

The samples were loaded onto a denaturing agarose gel. A 300 ml 1% gel was made by adding 3 g of agarose (SeaKem™ LE, FMC BioProducts, Rockland, ME) and 60 ml of 5x MOPS buffer to 210 ml sterile H2O. 5x MOPS buffer (0.1M MOPS (pH 7.0), 40 mM NaOAc, 5mM EDTA (pH 8.0)) was made by adding 20.6 g of MOPS to 800 ml of 50mM NaOAc (13.3 ml of 3M NaOAc pH 4.8 in 800 ml sterile H_2O); then adjusting the pH to 7.0 with 10M NaOH; adding 10 ml of 0.5M EDTA (pH8.0); and adding H_2O to a final volume of 1L. The mixture was heated until melted, then cooled to 50°C, at which time 5 μl ethidium bromide (5mg/ml) and 30 ml of 37% formaldehyde of gel were added. The gel was swirled quickly to mix, and then poured immediately.

2μg RNA sample, 1x final 1.5x RNA loading dyes (60% formamide, 9% formaldehyde, 1.5X MOPS, .075% XC/BPB dyes) and H_2O were mixed to a final volume of 40 μl. The tubes were heated at 65°C for 5 min. and then cooled on ice. 10 μg of RNA MW standards (New England Biolabs, Beverly, MA) were also denatured with dye and loaded onto the gel. The gel was run overnight at 32V in MOPS running buffer.

The gel was then soaked in 0.5 μg/ml Ethidium Bromide for 45 min., 50 mM NaOH/0.1 M NaCl for 30 min., 0.1 M Tris pH 8.0 for 30 min., and 20x SSC for 20 min. Each soaking step was done at r.t. with shaking. The gel was then photographed

along with a fluorescent ruler before blotting with Hybond-N membrane (Amersham), according to the methods of Sambrook et al., 1989, *supra*, in 20x SSC overnight.

For hybridization, the blot was placed into a roller bottle containing 10 ml of prehybridization solution consisting of 50% formamide and 1x Denhardt's solution, and placed into 65°C incubator for 30 min. The probe was then heated to 95°C, chilled on ice, and added to 10 ml of hybridization solution, consisting of 50% formamide, 1x Denhardt's solution, 10% dextransulfate, to a final concentration of 10^6 cpm/ml. The prehybridization solution was then replaced with the probe solution and incubated overnight at 42°C. The following day, the blot was washed three times for 30 min. in 2x SSC/0.1% SDS at room temperature before being covered in plastic wrap and put down for exposure.

RT-PCR Analysis: RT-PCR was performed to detect differentially expressed levels of mRNA from the genes corresponding to amplified bands. First strand synthesis was conducted by mixing 20 μl DNased RNA (~2 μg), 1 μl oligo dT (Operon) (1 μg), and 9.75 μl H_20. The samples were heated at 70°C for 10 min., and then allowed to cool on ice. 10 μl first strand buffer (Gibco/BRL), 5 μl 0.1M DTT, 1.25 μl 20 mM dNTP's (500 μM final), 1 μl RNAsin (40 units/μl) (Boehringer Mannheim), and 2 μl Superscript Reverse Transcriptase (200 units/μl) (Gibco/BRL) were added to the reaction, incubated at 42°C for 1 hr., and then placed at 85°C for 5 min., and stored at -20°C.

PCR was performed on the reverse transcribed samples. Each reaction contained 2 μl 10x PCR buffer, 14.5 μl H_20, 0.2 μl 20 mM dNTP's (200 μM final), 0.5 μl 20 μM forward primer (0.4 μM final), 0.5 μl 20 μM reverse primer (0.4 μM final), 0.3 μl AmpliTaq polymerase (Perkin-Elmer/Cetus), 2 μl cDNA dilution or positive control (~40 pg). Samples were placed in the PCR 9600 machine at 94°C (hot start), which was programmed as follows:

	94°C	2	min. (samples loaded)
	*94°C	45	sec.
* = 35x	*55°C	45	sec.
	*72°C	2	min.
	72°C	5	min.
	4°C hold		

Reactions were carried out on cDNA dilution series and tubes were removed at various cycles from the machine during 72°C step. Reaction products were eluted on a 1.8% agarose gel and visualized with ethidium bromide.

6.1.3. CHROMOSOMAL LOCALIZATION OF TARGET GENES

Once the nucleotide sequence has been determined, the presence of the gene on a particular chromosome is detected. Oligonucleotide primers based on the nucleotide sequence of the target gene are used in PCR reactions using individual human chromosomes as templates. Individual samples of each the twenty-three human chromosomes are commercially available (Coriel Institute for Medical Research, Camden, NJ). The chromosomal DNA is amplified according to the following conditions: 10ng chromosomal DNA, 2μl 10x PCR buffer, 1.6μl 2.5mM dNTP's, 0.1μl 25mM $MgCl_2$, 0.2μl reverse primer (100ng/μl), 0.2μl forward primer (100ng/μl), 0.1 μl Taq polymerase, and 15.8μl H_20. Samples are placed in the PCR 9600 machine at 94°C (hot start), which is programmed as follows:

	94°C	2	min. (samples loaded)
	*94°C	20	sec.
* = 35x	*55°C	30	sec.
	*72°C	30	sec.
	72°C	5	min.
	4°C hold		

7. EXAMPLE: IDENTIFICATION OF GENES DIFFERENTIALLY EXPRESSED IN RESPONSE TO PARADIGM B: IN VIVO MONOCYTES

In an alternative embodiment of the invention, genes differentially expressed in monocytes were detected under highly physiologically relevant, in vivo conditions. According to Paradigm B, human subjects were held in a clinical setting and the fat/cholesterol content of their diets was controlled. Monocytes were isolated at different stages of treatment, and their gene expression pattern was compared to that of control groups.

By use of Paradigm B, the human bcl-2 gene was identified. Its expression decreases in response to the atherogenic conditions of high fat/high cholesterol. The Apo E-/- mouse is the first mouse model of atherosclerosis with pathology similar to that of human plaque development (Plump et al., 1992, Cell 71: 343-353). Serum cholesterol levels in these mice on a chow diet is five times higher than those of control littermates. To address whether the regulation of the mouse bcl-2 gene is also affected by serum cholesterol levels, white blood cells, which include monocytes, from apoE-deficient mice and littermate wild-type controls were purified and mouse bcl-2 mRNA levels were compared using quantitative RT-PCR. By this method, mouse bcl-2 mRNA levels were significantly lower in the apoE-deficient mice relative to the wild-type controls.

The differential expression pattern of the human glutathione peroxidase gene (HUMGPXP1) was also discovered. The differential expression of HUMGPXP1 was initially detected in a preliminary detection system, described, below, in Section 7.1.2. Once HUMGPXP1 sequences were obtained, the gene's differential expression pattern was verified and characterized under the physiologically relevant conditions of Paradigm B. Glutathione peroxidase is known to be involved in the removal of toxic peroxides that form in the course of growth and metabolism under normal aerobic conditions and under oxidative stress. Human plasma glutathione peroxidase gene was originally isolated from a

human placenta cDNA library (Takahashi et al., 1990, J. Biochem. 108: 145-148). It has been shown to be expressed in two human cell lines of the myeloid lineage (Porter et al., 1992, Clinical Science 83: 343-345). Other studies have also linked reduced levels of this enzyme with heart attack risk (Guidi, et al., 1986, J. Clin. Lab Invest. 46: 549-551; Wang et al., 1981, Klin. Wochenschr. 59: 817-818; Kok et al., 1989, J. Am. Med. Assoc. 261: 1161-1164; and Gromadzinska & Sklodowska, 1990, J. Am. Med. Assoc. 263: 949-950). Glutathione peroxidase has not been previously known to be down-regulated in human monocytes under cardiovascular disease conditions, as described herein.

Interestingly, bcl-2 has been recognized as playing a key role in preventing apoptosis, and expression of glutathione peroxidase in the absence of bcl-2 is able to compensate for this loss by preventing apoptosis (Hockenbery et al., 1993, Cell 75: 241-251). These findings regarding bcl-2 and HUMGPXP1, described herein in this section, suggested a novel role for the monocyte in plaque formation which involves apoptosis induction caused by high LDL concentrations inside the cell, or perhaps by oxidative stress in the cell mediated by oxidized LDL.

To confirm this relationship between apoptosis and atherosclerosis, the ability of bcl-2 expression to ameliorate atherosclerosis is tested. Because bcl-2 is normally down-regulated under atherogenic conditions, a transgenic mouse strain is engineered in which the human bcl-2 gene is expressed under the control of the scavenger receptor promoter, which is induced in monocyte foam cells under atherogenic conditions. This transgenic mouse is then crossed with an apoE-deficient atherosclerotic mouse model. The ability of the increased expression of the bcl-2 target gene to ameliorate atherosclerosis is demonstrated by a decrease in initiation and progression of plaque formation observed in the transgenic apoE-deficient mouse.

The identification of the differential expression of these genes, therefore, provides targets for the treatment

and diagnosis of cardiovascular disease. Intervening in the apoptotic pathway through Bcl-2 and glutathione peroxidase, may lead to lesion regression or prevention of plaque formation, or both. Furthermore, the discovery of a connection between the apoptotic pathway and atherosclerosis demonstrates the effectiveness of the methods described herein in identifying the full panoply of gene products that are involved in the atherosclerotic disease process. Furthermore, the down-regulation of bcl-2 and HUMGPXP1 under Paradigm B provides a fingerprint for the study of the effect of excess LDL on monocytes.

7.1. MATERIALS AND METHODS

7.1.1. IN VIVO CHOLESTEROL STUDIES

Patients were held in a clinical setting for a total of 9 weeks during which time their lipid intake was very tightly controlled. There were a total of 3 diets, and each patient was held on each diet for 3 weeks. Patients were healthy young (third decade of life) individuals with no history or symptoms of heart disease or dislipidemias. The 3 diets are described below:

American Heart Association Diet II	
fat	25%
cholesterol	80 mg/1000 kCal
polyunsaturated/saturated fat	1.5
Average American Diet	
fat	43%
cholesterol	200 mg/1000 kCal
polyunsaturated/saturated fat	0.34
Combination Diet	
fat	43%
cholesterol	80 mg/1000 kCal
polyunsaturated/saturated fat	0.34

The 3 diets were isocaloric, and the individual components of each diet may vary with the participant's preference as long as the lipid levels in the diet were maintained.

Cell Isolation

At the end of each 3 week diet period, blood was drawn from each patient after a 12 hour period of fasting and monocytes were purified. 50 ml of blood was drawn into 5 evacuated tubes containing 1.4 ml each of citrate phosphate dextrose to prevent coagulation. Blood was pooled into 50 ml tubes and spun at 400g (1250 RPM/Sorvall RC3B) for 15 minutes at 4°C. The upper serum layer (~ 25 ml) was then removed with a pipette and replaced with phosphate buffered saline (PBS) at 4°C. The blood was mixed and then spun at 1850 x g (2680 RPM) for 15 minutes at 4°C. Most of the clear upper layer was removed with a pipette, before the buffy coat at the interface was taken in ~5 ml. The buffy coat was placed into a separate 50 ml tube, and the pipette used to remove it was washed with 20 ml PBS. A small aliquot of these cells was then diluted 1:1000 in PBS and counted under a microscope using a hemacytometer. Red blood cell concentration was then adjusted with PBS to a final concentration of 1.5×10^9/ml, and 10 ml aliquots were added to Leucoprep Becton Dickinson) tubes for monocyte isolation. Tubes were spun for 25 minutes at 25°C in a Sorvall RT6000 with the brake off. Most of the clear upper layer was discarded, and the turbid layer above the gel was saved and pooled in 50 ml tubes. The volume of each tube was then increased to 50 ml with 25°C PBS, and spun at 1000 RPM (Sorvall RC3B) for 10 minutes at 4°C. The liquid was then discarded, the pellet was resuspended in 50 ml PBS, and spun again. This process was repeated 3 more times. The final cell pellet was then resuspended in 2 ml RNA lysis buffer (Sambrook et al., 1989, supra) and frozen for subsequent RNA isolation as described above in Section 6.1.1.

Differential display, Northern analysis, RT-PCR, subcloning, and DNA sequencing were carried out as described, above, in Section 6.1.2.

7.1.2. PRELIMINARY DETECTION SYSTEM

The preliminary detection system described in this section was used to identify sequences that are differentially expressed in a readily assayed, in vitro system. Sequences that showed some homology to those thought to be involved in cardiovascular disease were then used as specific primers or probes, or both, in Paradigm B, wherein the differential expression was ascertained under physiologically relevant conditions, as described in section 7.1.1, above.

Cell culture Blood (~100 ml) was drawn from healthy human donors into vacutainer tubes containing heparin (Becton Dickinson). Blood was diluted 1:1 with PD (Phosphate buffered saline (PBS) without Ca or Mg, plus 0.3mM EDTA), and layered onto Ficoll (Lymphocyte Separation Media - Organon Teknikon) as 30 ml of blood/7 ml ficoll in a 50 ml blue-capped Falcon tube, and centrifuged at 2000 RPM for 25 min. at room temperature (r.t.). The buffy coat was removed with a pipette, transferred to another 50 ml tube, diluted to 30 ml with PD, and centrifuged at 1200 RPM for 10 min. at r.t. The pellet was resuspended in 30 ml PD and the previous centrifugation step was repeated. The pellet was resuspended in 40 ml RPMI (2mM 1-Glutamine + penicillin/streptomycin), plated onto 4 plates, and incubated at 37°C for 2 hours. Supernatant was removed, and the plates were washed 3x with PBS at 37°C. Plates were finally resuspended in 10 ml each with RPMI/20% human AB serum (Sigma, St. Louis, MO). On day 5, the media was changed and 100 units/ml of human γ-IFN (Genzyme) were added. On day 7, the media was removed and replaced with RPMI/20% human LDL-deficient serum + 100 units/ml of human γ-IFN. Native, oxidized, and acetylated LDL were each added to one plate with the fourth plate serving as control. After the specified incubation time (5 hr. or 24 hr.) the media was removed and the cells were resuspended in 2 ml guanidine isothiocyanate RNA lysis buffer (Sambrook et al., 1989, *supra*). Lysed cells were then syringed with 23 G. needle, layered over 5.7M CsCl, and

centrifuged for 20 hr. at 35K RPM. RNA was isolated according to the method of Sambrook et al., 1989, *supra*.

Lipoproteins were prepared as described, above, in section 6.1.1. Differential display, Northern analysis, RT-PCR, subcloning, and DNA sequencing were carried out as described, above, in Section 6.1.2. For differential display, the primers used were $T_{11}CC$ (reverse) and OPE4 (forward), consisting of 5'GTGACATGCC3'. For RT-PCR, the first strand cDNA was primed with $T_{11}CC$, and PCR reactions were carried out with rfhma15 primers (for-catgcctgtagaaaaaggtt/rev-cttcatagaatctaagccta), and mouse γactin primers (for-cctgatagatgggcactgtgt/rev-gaacacggcattgtcactaact).

7.1.3. TRANSGENIC ApoE-DEFICIENT MOUSE EXPRESSING HUMAN bcl-2

Transgenic mice bearing a construct with the mouse scavenger receptor regulatory element (5kb) (M. Freeman, et al., 1995, unpublished results) driving expression of the human bcl-2 gene (hbcl-2) were produced. The scavenger receptor regulatory element (ScR) is known to activate reporter gene expression in peritoneal macrophages in transgenic mice (M. Freeman, 1995, unpublished results). This 5 kb fragment is linked to the human bcl-2 cDNA (Cleary, et al., 1986, *supra*) via a NotI restriction site. Human growth hormone (hGH) sequences (Mayo, et al., 1983, Nature 306: 86-88) are then ligated onto the 3' end of this construct through filled-in BamHI and EcoRV sites to provide message stability. This construct is then digested with XhoI and the 9 kb ScR-hbcl2-hGH sequences are purified away from vector sequences. Another plasmid sample is digested with KpnI to yield a fragment with only 1.5kb of scavenger receptor regulatory sequences which provide a lower level of expression. These fragments are then injected independently into mouse embryos derived from the FVB and C57BL/6 mouse strains according to standard protocols (Hogan, et al., Manipulating the Mouse Embryo, 1994, Cold Spring Harbor

Laboratory Press). Following birth, tail sections are cut from mice derived from injected embryos and analyzed for the presence of transgene sequences using hbcl-2 sequences as probes on Southern blots.

Transgenic mice bearing the ScR-hbcl2-hGH construct are then bred to wild-type mice of the same respective strain, and then the offspring are backcrossed to produce homozygous lines of mice. These mice are then bred to apoE-deficient mice. Offspring are analyzed for presence of the ScR-hbcl2-hGH by preparing tail sections and probing with hbcl-2 sequences on Southern blots. Offspring are then analyzed for lesion formation and progression according to the methods of Plump, et al., 1992, *supra*.

7.2. RESULTS

Differential display analysis was carried out on monocyte RNA derived from the blood of patients whose serum cholesterol levels were manipulated through fat/cholesterol intake in their diets. A band designated band #14 which was present in the low dietary fat/low serum cholesterol conditions and goes away in the high dietary fat/high serum cholesterol conditions. When a radioactively labeled probe was prepared from band #14 and hybridized with a Northern blot prepared from RNA from the same patient, an 8 kb band was seen which was present in low serum cholesterol and disappeared in high serum cholesterol conditions. When band #14 sequences were subcloned, sequenced, and compared with the sequence database a 98% (203/207 bp) sequence similarity with the human bcl-2 gene (Cleary et al., 1986, Cell 47, 19-28) was obtained, indicating that band #14 is bcl-2.

Glutathione peroxidase (HUMGPXP1) in expression in monocytes was examined to determine its physiological relationship to bcl-2. Differential expression of HUMGPXP1 was first detected in a preliminary detection system using monocytes cultured in vitro. Human monocytes were prepared as described above in subsection 7.1.2. Cells were lysed after 5 hours and RNA was prepared. Differential display

analysis was carried out, and regulated bands were isolated and characterized. The DNA sequence was determined from a number of independent subclones of amplified sequences of one such regulated band designated band 15. Using the BLAST program (Altschul, et al., 1990, J. Mol. Biol. 215: 403-410), a 176/177 (99%) sequence similarity was found between band 15 a sequence for human plasma glutathione peroxidase exon 1 (HUMGPXP1). This sequence occurs upstream of the reported transcription start site. Nonetheless, RT-PCR analysis confirmed that the band 15 sequences are in fact within the same transcription unit as sequences downstream of the reported transcription start site.

Based on this preliminary result, the gene expression pattern of glutathione peroxidase (HUMGPXP1) was further analyzed for verification and characterization in physiologically relevant samples according to Paradigm B. Monocytes derived from human blood under atherogenic conditions (high serum cholesterol) and healthy conditions (low serum cholesterol) were examined with RT-PCR. There appeared to be 2-3 fold less cDNA amplified by the HUMGPXP1 primers from the high fat/cholesterol monocytes than in the low fat/cholesterol monocytes, while the actin control bands are the same.

White blood cells, which include monocytes, from apoE-deficient mice and littermate wild-type controls were purified and mouse bcl-2 mRNA levels were compared using quantitative RT-PCR. By this method, mouse bcl-2 mRNA levels were significantly lower in the apoE-deficient mice relative to the wild-type controls.

These results demonstrate that bcl-2 is an excellent target gene for intervening in lesion formation and development. It was previously known that, under normal conditions, bcl-2 expression prevents apoptosis. The observed down-regulation of bcl-2 caused by atherogenic conditions, therefore, provides an explanation of how such atherogenic conditions may lead to plaque formation. By down-regulating the normally protective bcl-2 gene, high

serum cholesterol triggers a series of events, entailing the induction of the apoptotic pathway, which results in programmed cell death, which in turn causes an inflammatory response and subsequent plaque formation.

This model may be tested by counteracting the observed down-regulation of bcl-2. The human bcl-2 gene is placed in the ScR-hbcl2-hGH construct in which it is transcribed by a promoter that is activated in monocyte foam cells under atherogenic conditions. This construct is then introduced into an apoE-deficient mouse that otherwise serves as a model for atherosclerosis. The effect of bcl-2 expression on atherosclerosis is evidenced by the reduction in plaque initiation and development in the apoE-deficient mice bearing the construct. Amelioration of atherosclerosis may, therefore, be accomplished by such intervention in the down-regulation of the bcl-2 target gene.

8. EXAMPLE: IDENTIFICATION OF GENES DIFFERENTIALLY EXPRESSED IN RESPONSE TO PARADIGM C: IL-1 INDUCTION OF ENDOTHELIAL CELLS

According to the invention, differential display was used to detect four novel genes that are differentially expressed in endothelial cells that were treated in vitro with IL-1. Three of these genes, rchd024, rchd032, and rchd036, are not homologous to any known gene. The fourth gene, rchd005, is 70% homologous to a cloned shark gene called bumetanide-sensitive Na-K-Cl cotransport protein. A human homolog of this gene has been reported, but the sequence has not yet been published (Xu et al., 1994, Proc. Natl. Acad. Sci. USA 91: 2201-2205).

The discovery of the up-regulation of these four genes provides a fingerprint profile of IL-1 induced endothelial cells. This fingerprint profile can be used in the treatment and diagnosis of cardiovascular diseases, including but not limited to atherosclerosis, ischemia/reperfusion, hypertension, restenosis, and arterial inflammation.

8.1. MATERIALS AND METHODS

Primary cultures of HUVEC's were established from normal term umbilical cords as described (In Progress in Hemostasis and Thrombosis, Vol. 3, P. Spaet, editor, Grune & Stratton Inc., New York, 1-28). Cells were grown in 20% fetal calf serum complete media (Luscinskas, et al., 1989, J. Immunol. 142: 2257-2263) and passaged 1-3 times before activation.

For activation, cells were cultured with 10 units/ml of human IL-1β for 1 or 6 hr. before lysis in guanidinium isothiocyanate RNA lysis buffer (Sambrook et al., 1989, supra). Lysed cells were then syringed with a 23 G. needle, layered over 5.7M CsCl, and centrifuged for 20 hr. at 35K.

Alternatively, cells were induced in the presence of 100μM lysophosphatidylcholine, or 50 μg/ml oxidized human LDL (Sigma) for periods of 1 or 6 hr. RNA was isolated as described, above, in Section 6.1. Differential display, Northern analysis, RT-PCR, subcloning, and DNA sequencing were carried out as described, above, in Section 6.1.2, except that Northern blot hybridizations were carried out as follows: for pre-hybridization, the blot was placed into roller bottle containing 10 ml of rapid-hyb solution (Amersham), and placed into 65°C incubator for at least 1 hr. For hybridization, $1x10^7$ cpm of the probe was then heated to 95°C, chilled on ice, and added to 10 ml of rapid-hyb solution. The prehybridization solution was then replaced with probe solution and incubated for 3 hr at 65°C. The following day, the blot was washed once for 20 min. at r.t. in 2x SSC/0.1% SDS and twice for 15 min. at 65°C in 0.1x SSC/0.1% SDS before being covered in plastic wrap and put down for exposure.

Chromosomal locations were determined according to the method described in Section 6.1.3, above. For rchd024, the primers used were for-cccatagactaggctcatag, and rev-tttaaagagaaattcaaatc.

8.2. RESULTS

HUVEC's were activated with 10 units/ml IL-1β for 1 or 6 hours and compared to resting HUVEC's using differential display. A band designated rchd005 was present in lanes 11 and 12 (IL-1, 6 hr.) but not in lanes 9 and 10 (control), or lanes 7 and 8 (IL-1, 1 hr.). This band, rchd005, was isolated and subcloned and sequenced. When a probe prepared form this band was used to screen a Northern blot, expression was seen at 6 hr., but not at 1 hr. or in the control. However, when this same probe was hybridized to a Northern blot prepared from shear stressed RNA, according to Paradigm D described in Section 9, below, a different pattern of up-regulation was also seen. Expression was up at 1 hr. and then nearly disappeared by 6 hr. Amplified rchd005 DNA was subcloned and sequenced. Sequence analysis revealed an approximately 360 bp insert (FIG.1) with 70% sequence similarity to a cloned shark gene called bumetanide-sensitive Na-K-Cl cotransport protein.

Another IL-1 inducible differential display band was designated rchd024. Northern analysis on IL-1 up-regulated RNA revealed a 10 kb rchd024 message present at 6 hr. that also showed a low level of up-regulation under shear stress at 6 hr. The DNA sequence of rchd024 was obtained from subclones of amplified DNA (FIG.2). Database searching revealed no significant sequence similarities. A PCR amplification experiment determined that the rchd024 gene is located on human chromosome 4.

Band rchd032 was isolated on the basis of its differentially increased expression after 6 hr. treatment with IL-1, which was confirmed by RT-PCR analysis. Amplified rchd032 sequences were subcloned and sequenced (FIG.3). No significant homology to any known gene was found.

Band rchd036 was also isolated on the basis of its differential expression 6 hr. after IL-1 treatment. Northern analysis revealed an 8 kb band which was up-regulated 6 hr. after IL-1 treatment. Another Northern analysis was performed testing rchd036 under the shear stress condition of

Paradigm D, which are described in the example in Section 9, below. Interestingly, rchd036 is not induced by shear stress, as indicated by the lack of any band after either 1 hr. or 6 hr. of treatment. This result provides an example of an IL-1-inducible endothelial cell gene that is not regulated by shear stress, indicating that these induction pathways can be separated, and may provide for drugs with greater specificity for the treatment of inflammation and atherosclerosis. The DNA sequence was obtained from subclones of amplified DNA (FIG.4), and a search of the database revealed no sequence similarities. A PCR amplification experiment determined that the rchd036 gene is located on human chromosome 15.

9. EXAMPLE: IDENTIFICATION OF GENES DIFFERENTIALLY EXPRESSED IN RESPONSE TO PARADIGM D: ENDOTHELIAL CELL SHEAR STRESS

According to the invention, differential display was used to detect genes that are differentially expressed in endothelial cells that were subjected to fluid shear stress in vitro. Shear stress is thought to be responsible for the prevalence of atherosclerotic lesions in areas of unusual circulatory flow. Using the method of Paradigm D, four bands with novel DNA sequences were identified.

rchd502 is homologous to rat matrin F/G mRNA sequence (Hakes, et al., 1991, Proc. Natl. Acad. Sci. U.S.A. 88:6186-6190). This rat gene has been shown to encode a protein which functions as a prostaglandin transporter, and has been designated PGT (Kanai et al., 1995, Science 268: 866-869). In fact, the sequences in rchd502 encode the homologous twelve transmembrane domains found in the PGT gene. Furthermore, rchd502 was demonstrated to be up-regulated by shear-stress but not by IL-1. It therefore provides an excellent novel tool for diagnosis and treatment of cardiovascular disease.

The complete sequence of the rchd523 gene reveals that it encodes a novel G protein-coupled receptor protein,

consisting of 375 amino acids and seven transmembrane domains. At the amino acid level, rchd523 is 40% indentical to the Angiotensin II receptor. The discovery of such a novel protein is particularly useful in designing treatments as well as diagnostic and monitoring systems for cardiovascular disease. In carrying out signal transduction, G proteins play an important early role in the pathways that cause changes in cellular physiology. The rchd523 gene product, therefore, provides an excellent target for intervention in the treatment of cardiovascular disease.

The sequence of the coding region for rchd528 was partially determined. Sequence alignment revealed that the partial rchd528 sequence contains an extracellular domain with particularly strong homology to epidermal growth factor (EGF) repeats.

Furthermore, as transmembrane proteins, the rchd502, rchd523, and rchd528 gene products can be readily accessed or detected on the endothelial cell surface by other compounds. They provide, therefore, excellent targets for detection of cardiovascular disease states in diagnostic systems, as well as in the monitoring of the efficacy of compounds in clinical trials. Furthermore, the extracellular domains of these four gene products provide especially efficient screening systems for identifying compounds that bind to them. Such compounds, can be useful in treating cardiovascular disease by modulating the activity of the transmembrane gene products.

The sequence of the complete coding region of the rchd534 gene was also obtained. The rchd534 gene encodes a novel protein consisting of 235 amino acids, homologous to the MH2 domain of *Drosophila* protein *Mothers against decapentaplegic* (*Mad*) (Sekelsky et al., 1995, Genetics 139: 1347-1358; Hoodless, et al., 1996, Cell 85: 489-500). The rchd534 gene is also significantly similar to a sequence of unknown function from *Caenorhabditis elegans*, identified in the *C. elegans* genome project (Wilson, et al., 1994, Nature 368: 32-38). MAD is in the same pathway as Decapentaplegic (dpp), which is a Drosophila homolog of bone morphogenic

protein-4/Transforming growth factor-β (TGF-β). As described in detail in Section 16, below, the 235 amino acid rchd534 protein is encoded by the shorter of two spliceoforms of the rchd534 gene.

Also using the method of Paradigm D, the previously identified human prostaglandin endoperoxide synthase type II, also known as cyclooxygenase II (COX II), was identified (band rchd505). This gene was previously known to be involved in inflammation, and to be up-regulated by IL-1 (Jones et al., 1993, J. Biol. Chem. 268: 9049-9054), but its up-regulation by shear stress was previously unknown. This result confirmed the general effectiveness of the techniques used according to the invention in the detection of genes involved cardiovascular disease.

The sequence of another up-regulated gene, designated as rchd530, was shown to be identical to the previously identified human manganese superoxide dismutase gene (MnSOD). The up-regulation of MnSOD under shear stress was not previously known.

The up-regulation of these six genes in shear stressed endothelial cells provides a fingerprint for the study of cardiovascular diseases, including but not limited to atherosclerosis, ischemia/reperfusion, hypertension, and restenosis. The fact that one of these genes, rchd502, is not up-regulated under Paradigm C (IL-1 induction) provides an extremely useful means of distinguishing and targeting physiological phenomena specific to shear stress.

The importance of the induction of these genes in endothelial cells under disease conditions was further analyzed by testing the effect of estrogen on their expression. Studies in postmenopausal women on estrogen replacement therapy and in animal models have demonstrated that estrogen has an atheroprotective effect in reducing incidence of coronary artery disease (Gura, T., 1995, Science 269:771-773). While these studies demonstrate that estrogen has an effect in the liver in reducing LDL levels and increasing HDL levels, these lipoprotein changes are not

thought to be responsible for all of the cardioprotective effects of estrogen.

The identification of target genes that are differentially expressed under certain disease conditions provides for further analysis of the effect of estrogen on cardiovascular disease. The effect of estrogen on target gene expression in endothelial cells was, therefore, compared to particular paradigm expression patterns. Specifically, given that the estrogen receptor is a transcription factor (Kumar and Chambon, Cell 55:145-156, 1988), genes that are induced by shear stress were examined for regulation by estrogen in HUVEC's. In addition to estrogen, treatment with estrogen receptor agonists/antagonists tamoxifen (Grainger et al., Nature Medicine 1:1067-1073, 1995) and raloxifene (Black et al., J. Clin. Invest. 93:63-69, 1994), which also have been reported to have cardioprotective effects, were examined. The results demonstrate that rchd528, which is up-regulated by shear stress, is also up-regulated by estrogen, and suggest that shear stress and estrogen may play similar roles in cardiovascular disease.

9.1. MATERIALS AND METHODS

Primary cultures of HUVEC's were established from normal term umbilical cords as described (In Progress in Hemostasis and Thrombosis, Vol. 3, P. Spaet, editor, Grune & Stratton Inc., New York, 1-28). Cells were grown in 20% fetal calf serum complete media (Luscinskas et al., 1989, J. Immunol. 142: 2257-2263) and passaged 1-3 times before shear stress induction.

For induction, second passage HUVEC's were plated on tissue culture-treated polystyrene and subjected to 10 dyn/cm^2 laminar flow for 1 and 6 hr. as described (Nagel et al., 1994, J. Clin. Invest. 94: 885-891) or 3-10 dyn/cm^2 turbulent flow as previously described (Davies et al., 1986 Proc. Natl. Acad. Sci. U.S.A. 83: 2114-2117).

To examine the effect of estrogen on target gene expression, HUVEC's were cultured in serum free endothelial

cell basal medium supplemented with 1ug/ml insulin, 1ug/ml transferrin, 50 ug/ml gentamycin, and 200ug/ml fatty acid-free BSA. Cells were treated with either estradiol, tamoxifen, or raloxifene at 1nm final concentration for 4 or 16 hours before lysis and RNA isolation. For rchd528, the DNA fragment comprising bases 1600-2600 was used as a probe in Northern analysis.

RNA was isolated as described, above, in Section 6.1. Differential display, Northern analysis, RT-PCR, subcloning, and DNA sequencing were carried out as described, above, in Section 6.1.2, except that Northern blot hybridizations were carried out as described, above, in Section 8.1.

cDNAs containing larger portions or complete coding regions of the genes were obtained either by RACE, or by probing cDNA libraries, or both. The RACE procedure was carried out using a kit according to the manufacturer's instructions (Clontech, Palo Alto, CA; see also: Chenchik, et al., 1995, CLONTECHniques (X) 1: 5-8; Barnes, 1994, Proc. Natl. Acad. Sci. USA 91: 2216-2220; and Cheng et al., Proc. Natl. Acad. Sci. USA 91: 5695-5699). Primers were designed based either on amplified sequences, or on sequences obtained from isolates from the cDNA libraries. Template mRNA was isolated from shear stressed HUVEC's.

Amplified sequences, which contained portions of the genes, were subcloned and then used individually to retrieve cDNAs encoding the corresponding gene within cDNA libraries. Probes were prepared by isolating the subcloned insert DNA from vector DNA and labeling with ^{32}P as described above in Section 6.1.2. The libraries used included individual human heart, human pancreas, and human lung cDNA libraries, (Clontech, Palo Alto, CA); and a cDNA library prepared from mRNA which was isolated from shear stressed HUVEC's as described in this section, above. The HUVEC cDNA library was produced according to well-known methods (Sambrook et al., 1989, *supra*), using the bacteriophage λ-ZAP vector (Stratagene, La Jolla, CA). Libraries were screened by each respective probe using well-known methods (Sambrook et al.,

1989, *supra*). Plaques from the libraries that were detected by the probes were isolated and the cDNA insert within the phage vector was sequenced.

Determination of chromosomal location was carried out according to the method described in Section 6.1.3, above. The primers used for rchd523 were (for-atgccgtgtgggttagtc) and (rev-attttatgggaaggtttttaca); and for rchd534 were (for-cttttctgcgtctcccat) and (rev-agacatcagaaactccaacc).

Northern blot analyis of RNA extracted from various human organs and tissues was performed using commercially available pre-blotted filters (Clontech, Palo Alto, CA).

9.2. RESULTS

HUVEC's were subjected to laminar shear stress for 1 or 6 hr. and compared to static control cells in differential display. A band (rchd502) was identified which was found in lanes 5,6 (6 hr.) but not in lanes 1,2 (control). This band was excised, amplified, and sequenced. Northern analysis using amplified rchd502 sequences revealed a 4.5 kb band that is up-regulated at 6 hr. compared to controls. When rchd502 probe was hybridized to a Northern blot prepared from IL-1 induced endothelial cells, up-regulation of a 4.5 kb band was not seen. This result provides the first example of a shear stress-inducible endothelial cell gene that is not regulated by IL-1, indicating that these induction pathways can be separated, and may provide for drugs with greater specificity for the treatment of inflammation and atherosclerosis. The sequence of the amplified region of rchd502 was used to design probes for cloning the entire gene.

Both 5' and 3' RACE reactions were carried out to obtain a 2.2kb cDNA containing the entire coding sequence of the rchd502 gene. Based on the sequence information from RACE, a phage clone was isolated from a human pancreas library which contains all but the first 200 base pairs of the rchd502 coding region. This clone was designated pFCHD502SF. The remaining 200 base pairs were obtained through amplification from a human lung library by PCR with

specific primers. A fragment comprising base pairs 1-265 of the rchd502 gene was subcloned into the TA cloning vector to produce plasmid pFCHD502SJ. Thus, rchd502 is represented by two subclones, pFCHD502SJ comprising base pairs 1-265, and pFCHD502SF comprising base pairs 201 through the 3' end of the coding region, including 3' untranslated sequence.

The complete sequence encompassing the entire coding region is shown in FIG.5. rchd502 shows strong homology (81.4%) to the rat PGT gene, which encodes a prostaglandin transporter (Kanai et al., 1995, *supra*). It contains twelve transmembrane (TM) domains. The approximate bounds of each of the twelve TM domains are as follows:

TM1: about amino acid 31 to about amino acid 52.
TM2: about amino acid 68 to about amino acid 89.
TM3: about amino acid 102 to about amino acid 121.
TM4: about amino acid 173 to about amino acid 194.
TM5: about amino acid 206 to about amino acid 227.
TM6: about amino acid 259 to about amino acid 280.
TM7: about amino acid 315 to about amino acid 337.
TM8: about amino acid 366 to about amino acid 385.
TM9: about amino acid 403 to about amino acid 423.
TM10: about amino acid 510 to about amino acid 530.
TM11: about amino acid 555 to about amino acid 575.
TM12: about amino acid 607 to about amino acid 627.

Shear stress band rchd505 decreased 1 hr. and 6 hr. after shear stress, as compared to untreated control cells. Northern analysis revealed differential expression except that rchd505 was up-regulated after 1 hr. and 6 hr. shear stress treatment. This same band was similarly up-regulated in cells treated with IL-1 according to Paradigm C. Sequence analysis revealed that rchd505 is the previously characterized human endoperoxide synthase type II, also known as cyclooxygenase II (COX II).

rchd523 was detected under differential display as a band up-regulated after 1 hr. and 6 hr. shear stress treatment. The 6 hr. up-regulation of rchd523 was confirmed by RT-PCR. Amplified rchd523 sequences were subcloned, and

an isolate was sequenced and designated pRCHD523. The RACE procedure was used to obtain a 2.5 kb cDNA containing the entire coding sequence of the rchd523 gene. The cDNA isolate containing the complete coding sequence of rchd523 is designated pFCHD523. The DNA sequence comprising the complete coding region of the rchd523 gene is shown in FIG.6. Sequence analysis revealed that the rchd523 gene product encodes a novel G protein-coupled receptor, consisting of 375 amino acids and seven transmembrane domains. At the amino acid level, rchd523 is 40% indentical to the Angiotensin II receptor. A PCR amplification experiment determined that the rchd523 gene is located on human chromosome 7.

rchd528 was also detected as an up-regulated band after 1 hr. and 6 hr. shear stress treatment. This result was confirmed by Northern analysis in which probes of rchd528 amplified sequence detected an approximately 8 kb message that was up-regulated moderately after 1 hr., and up-regulated very strongly after 6 hr. The amplified sequences were subcloned and sequenced. This sequence information was used for initial probing of a cDNA library to isolate the rchd528 gene.

The amplified sequence was used for initial probing of a shear stressed HUVEC cDNA library to isolate a partial clone of rchd528. The RACE procedure was then used in combination with probing a human heart cDNA library and PCR amplification to obtain overlapping clones encompassing the entire rchd528 coding region. The complete coding region of the rchd528 gene is contained in the following three plasmids each containing a segment of the rchd528 gene cloned into pBluescript: pFCHD528A, comprising nucleotides 1-1200; pFCHD528B, comprising nucleotides 237-2982; and pFCHD528C, comprising nucleotides 2982 through the 3' end of the coding region. The DNA sequence comprising the complete coding region of the rchd528 gene is shown in FIG.7.

Based on homology to a number of different proteins, the rchd528 gene product was shown to contain an extracellular domain comprising the epidermal growth factor

(EGF) repeat motif. The approximate bounds of the EGF repeat are from about amino acid 1089 to about amino acid 1122. There is a signal peptide domain extending from about amino acid 5 to about amino acid 28. Also, there is a transmembrane domain extending from about amino acid 1348 to about amino acid 1370. In addition, there is an asparagine hydroxylation site consensus sequence from about amino acid 1140 to about amino acid 1151. Northern blot analysis of mRNA isolated from a variety of human organs and tissues revealed that rchd528 is very highly expressed in the heart.

The effect of estrogen on the expression of rchd528 in endothelial cells was also examined. Northern blot analysis revealed significant up-regulation of rchd528 after overnight treatment with estrogen compared to control cells.

A band designated rchd530 corresponded to a sequence strongly up-regulated in HUVECs after six hours of shear stress. This up-regulation is greater for laminar shear stress than for turbulent shear stress. Sequence analysis revealed that rchd530 is identical to human manganese superoxide dismutase (MnSOD). The induction of MnSOD by shear stress was not previously known. MnSOD was also demonstrated to be induced by six hours of treatment with IL-1.

rchd534 also was detected as being up-regulated in response to shear stress. Northern analysis revealed that rchd534 is strongly induced after 6 hours of shear stress treatment (FIG.12). The amplified sequences were subcloned, sequenced, and re-isolated for use as a probe for retrieving full-length rchd534 cDNA. A 3.3kb λ-ZAP clone was sequenced to reveal a full-length rchd534 cDNA (FIG.8). This clone containing the entire coding region the rchd534 protein was designated pFCHD534. The encoded protein consists of 235 amino acids. A PCR amplification experiment determined that the rchd534 gene is located on human chromosome 15.

An initial comparison with sequences in the database revealed no homologies between rchd534 and any known DNA sequences. A subsequently performed search revealed that

rchd534 is a homolog of the *Drosophila* gene *Mothers against decapentaplegic (Mad)* (Sekelsky et al., 1995, Genetics 139: 1347-1358), and is also significantly similar to a sequence of unknown function from *Caenorhabditis elegans*, identified in the *C. elegans* genome project (Wilson, et al., 1994, Nature 368: 32-38). The 235 amino acid rchd534 protein contains an MH2 ("Mad homology") domain (see Hoodless, et al., 1996, Cell 85: 489-500 for description of Mad homology domains MH1 and MH2). As detailed in Section 16, below, the 235 amino acid rchd534 protein is encoded by the shorter of two spliceoforms of the rchd534 gene. The discovery of a longer spliceoform, and the protein it encodes designated rchd534-long, is described in detail in Section 16, below.

The expression of rchd534 was also shown not to be regulated by IL-1 when tested under the conditions of Paradigm C, as described in Section 8, above. Just like rchd502, rchd534 is an example of a shear stress-inducible endothelial cell gene that is not regulated by IL-1, confirming that these induction pathways can be separated, and may provide for drugs with greater specificity for the treatment of inflammation and atherosclerosis.

10. EXAMPLE: USE OF GENES UNDER PARADIGM A AS SURROGATE MARKERS IN CLINICAL TRIALS

According to the invention, the fingerprint profile derived from any of the paradigms described in Sections 5.1.1.1 through 5.1.1.6 may be used to monitor clinical trials of drugs in human patients. The fingerprint profile, described generally in Section 5.5.4, above, indicates the characteristic pattern of differential gene regulation corresponding to a particular disease state. Paradigm A, described in Section 5.1.1.1, and illustrated in the example in Section 6, above, for example, provides the fingerprint profile of monocytes under oxidative stress. This profile gives an indicative reading, therefore, of the physiological response of monocytes to the uptake of oxidized LDL. Accordingly, the influence of anti-oxidant drugs on the

oxidative potential may be measured by performing differential display on the monocytes of patients undergoing clinical tests.

10.1. TREATMENT OF PATIENTS AND CELL ISOLATION

Test patients may be administered compounds suspected of having anti-oxidant activity. Control patients may be given a placebo.

Blood may be drawn from each patient after a 12 hour period of fasting and monocytes may be purified as described, above, in Section 7.1.1. RNA may be isolated as described in Section 6.1.1, above.

10.2. ANALYSIS OF SAMPLES

RNA may be subjected to differential display analysis as described in Section 6.1.2, above. A decrease in the physiological response state of the monocytes is indicated by a decreased intensity of those bands that were up-regulated by oxidized LDL under Paradigm A, and an increased intensity of those bands that were down-regulated by oxidized LDL under Paradigm A, as described in Section 6.2, above.

11. EXAMPLE: IMAGING OF A CARDIOVASCULAR DISEASE CONDITION

According to the invention, differentially expressed gene products which are localized on the surface of affected tissue may be used as markers for imaging the diseased or damaged tissue. Conjugated antibodies that are specific to the differentially expressed gene product may be administered to a patient or a test animal intravenously. This method provides the advantage of allowing the diseased or damaged tissue to be visualized non-invasively.

For the purposes of illustration, this method is described in detail for the rchd523 gene product. The principles and techniques can be applied to any identified transmembrane target gene product, including, for example, the rchd502 and rchd528 gene products.

11.1. MONOCLONAL CONJUGATED ANTIBODIES

The differentially expressed surface gene product, such as the rchd523 gene product, is expressed in a recombinant host and purified using methods described in Section 5.4.2, above. Preferably, a protein fragment comprising one or more of the extracellular domains of the rchd523 product is produced. Once purified, it is be used to produce $F(ab')_2$ or Fab fragments, as described in Section 5.4.3, above. These fragments are then labelled with technetium-99m (^{99m}Tc) using a conjugated metal chelator, such as DTPA as described in section 5.8.3, above.

11.2. ADMINISTRATION AND DETECTION OF IMAGING AGENTS

Labeled MAb may be administered intravenously to a patient being diagnosed for atherosclerosis, restenosis, or ischemia/reperfusion. Sufficient time is allowed for the detectably-labeled antibody to localize at the diseased or damaged tissue site (or sites), and bind to the rchd523 gene product. The signal generated by the label is detected by a photoscanning device. The detected signal is then converted to an image of the tissue, revealing cells, such as endothelial cells, in which rchd523 gene expression is up-regulated.

12. EXAMPLE: SCREENING FOR LIGANDS OF THE rchd523 GENE PRODUCT AND ANTAGONISTS OF rchd523 GENE PRODUCT-LIGAND INTERACTION

The rchd523 gene product is a member of the G protein-coupled receptor protein family, containing multiple transmembrane domains. The receptor binding activity of this protein family is detected by assaying for Ca^{2+} mobility through the membrane of cells in which the receptor gene is expressed. This assay, described below, is used to identify ligands that bind to the rchd523 gene product receptor. Establishing this ligand-receptor activity then provides for a screen in which antagonists of the ligand-receptor interaction are identified. An antagonist is detected by its

ability to inhibit the Ca^{2+} mobility induced by ligand-receptor binding. Such antagonists, therefore, provide compounds that are useful in the treatment of cardiovascular disease, by counteracting the activity of the product of this target gene which is up-regulated in the disease state.

Binding of ligand to the rchd523 gene product is measured as follows. The cDNA containing the entire coding region of the rchd523 gene is removed from pFCHD523 and placed under the control of a promoter that is highly expressed in mammalian cells in an appropriate expression vector. The resulting construct is transfected into myeloma cells, which are then loaded with FURA-2 or INDO-1 by standard techniques. Ligands are added to the cell culture to test their ability to bind to the rchd523 receptor in a manner that triggers signal transduction, as measured by Ca^{2+} mobilization across the cell membrane. Mobilization of Ca^{2+} induced by ligand is measured by fluorescence spectroscopy as described in Grynkiewicz et al., 1985, *J. Biol. Chem.* 260:3440. Ligands that react with the target gene product receptor domain are identified by their ability to produce a fluorescent signal. Their receptor binding activities are quantified and compared by measuring the level of fluorescence produced over background.

Candidate antagonists are then screened for their ability to interfere with ligand-receptor binding. Myeloma transfectants expressing rchd523 gene product are treated with ligand alone, and ligand in the presence of candidate antagonist. Candidate antagonists that cause a reduction in the fluorescence signal are designated antagonists of the ligand-rchd523 receptor interaction.

13. POLYCLONAL ANTIBODIES TO TARGET GENE PEPTIDE SEQUENCES

Peptide sequences corresponding to the indicated amino sequences of cDNAs were selected and submitted to Research Genetics (Huntsville, AL) for synthesis and antibody production. Peptides were modified as described (Tam, J.P., 1988, Proc. Natl. Acad. Sci. USA 85: 5409-5413; Tam, J.P.,

and Zavala, F., 1989, J. Immunol. Methods 124: 53-61; Tam, J.P., and Lu, Y.A., 1989, Proc. Natl. Acad. Sci. USA 86: 9084-9088), emulsified in an equal volume of Freund's adjuvant and injected into rabbits at 3 to 4 subcutaneous dorsal sites for a total volume of 1.0 ml (0.5 mg peptide) per immunization. The animals were boosted after 2 and 6 weeks and bled at weeks 4, 8, and 10. The blood was allowed to clot and serum was collected by centrifugation.

The peptides used are summarized below:

rchd502 Protein

Peptide	Amino Acids #'s	Sequence
fchd502.1	294-308	DEARKLEEAKSRGSL
fchd502.2	435-449	SSIHPQSPACRRDCS
fchd502.3	627-640	RVKKNKEYNVQKAA

rchd523 Protein

fchd523.1	243-258	RAHRHRGLRPRRQKAL
fchd523.2	360-372	IPDSTEQSDVRFS

rchd528 Protein

fchd528.1	1393-1410	SPYAEYPKNPRSQEWGRE
fchd528.2	1467-1481	NPSFISDESRRRDYF

rchd534 and rchd534-long Proteins

	rchd534	rchd534-long	
fchd534.1	54-69	-	EFSDASMSPDATKPSH
fchd534.2	112-125	373-386	LEQRSESVRRTRSK
fchd534.3	182-197	443-458	RSGLQHAPEPDAADGP

14. LOCALIZATION OF NOVEL GENES BY *IN SITU* HYBRIDIZATION

The expression of two target genes, rchd502 and rchd528, was examined by *in situ* hybridization. The expression was detected in human carotidendarterectomy samples, i.e., human cardiovascular tissue in a diseased state, taken from a living patient suffering from cardiovascular disease. The expression pattern for each gene was observed to be similar to the pattern detected for the positive control, which is known to be constitutively

expressed in endothelial cells. These results provide further evidence of the role of both rchd502 and rchd528 in cardiovascular disease. The detection of high levels of expression of these target genes specifically within the endothelial cells of diseased tissues allows for more precise diagnosis, as well as more precise treatment methods, than simple detection of atherosclerotic lesion provides.

14.1 Methods

7μm paraffin embedded sections of human carotid endarterectomy samples were deparaffinized in xylenes, rehydrated through graded ethanol series and post-fixed with 4% PFA/PBS for 15 minutes. After washing with PBS, sections were digested with 2 μg/ml proteinase K at 37° for 15 minutes, and again incubated with 4% PFA/PBS for 10 minutes. Sections were then washed with PBS, incubated with 0.2 N HCl for 10 minutes, washed with PBS, incubated with 0.25% acetic anhydride/1 M triethanolamine for 10 minutes, washed with PBS and dehydrated with 70% ethanol and 100% ethanol.

Hybridizations were performed with ^{35}S-radiolabeled ($5x10^7$ cpm/ml) cRNA probes encoding 1) the 0.8 kB SmaI fragment segment of the coding region of the human von Willebrand factor gene, 2) a fragment containing portions of the novel gene rchd502 (sequence base pairs 3-1195, excluding bases 396-622), and 3) a fragment of the novel gene fchd528 (sequence base pairs 3718-6407) in the presence of 50% formamide, 10% dextran sulfate, 1x Denhardt's solution, 600 mM NaCl, 10 mM DTT, 0.25% SDS and 100 μmg/ml tRNA for 18 hours at 55°. After hybridization, slides were washed with 5x SSC at 55°, 50% formamide/2x SSC at 55° for 30 minutes, 10 mM Tris-HCl(pH 7.6)/500 mM NaCl/1 mM EDTA (TNE) at 37° for 10 minutes, incubated in 10 μg/ml RNase A in TNE at 37° for 30 minutes, washed in TNE at 37° for 10 minutes, incubated once in 2x SSC at 50° for 30 minutes, and dehydrated with 70% ethanol and 100% ethanol. Localization of mRNA transcripts was detected by dipping slides in Kodak NBT-2 photoemulsion and exposing for 7 days at 4°, followed by development with

Kodak Dektol developer. Slides were counterstained with Haemotoxylin and Eosin and photographed. Controls for the *in situ* hybridization experiments included the use of a sense probe which showed no signal above background levels.

14.2 Results

The rchd502 and rchd528 genes each displayed a similar expression patter to the positive control signal from von Willebrand factor, a constitutively expressed endothelial cell marker. Signal was detected for both rchd502 and rchd528 over most endothelial cells lining the luminal surface of the carotid artery, as was also observed for von Willebrand factor. None of the three genes examined showed expression in any other cell type present in the tissue, including smooth muscle cells and macrophages.

15. EXAMPLE: THE RCHD534 AND FCHD540 GENE PRODUCTS INTERACT

The amino acid sequence of the novel rchd534 protein, and the cDNA that encodes it, are described in detail Section 9, above. The novel rchd534 gene shares homology with the *Drosophila* Mad gene. The novel fchd540 gene (described in co-pending Application No. 08/799,910, filed February 13, 1997, which is incorporated by reference in its entirety herein) is another gene that is up-regulated in endothelial cells by shear stress. The DNA and encoded amino acid sequence of the fchd540 gene is shown in FIG.10. The fchd540 gene was deposited in the American Type Culture Collection (ATCC) in microorganism pFCHD540 on February 7, 1996 and assigned the ATCC Accession No. 69984. The fchd540 gene also shares homology with the *Drosophila* Mad gene. The fchd540 protein has both an MH1 domain and an MH2 domain. Mad genes have been shown to play a role in the TGF-β signalling pathway (Sekelsky et al., 1995, Genetics 139: 1347-1358; Chen et al., 1996, Nature 383: 691-696; Serra, et al., 1996, Nature Medicine 2: 390-391). TGF-β signalling is considered to be beneficial to atherosclerosis and restenosis (Border et al., 1995, Nature Medicine 1: 1000; Grainger, et al., 1995,

Nature Medicine 1: 1067-1073; Kojima, et al., 1991, J. Cell Biol. 113: 1439-1445; Nikol, et al., 1992, J. Clin. Invest. 90: 1582-1592).

The data described below demonstrate that the rchd534 and fchd540 proteins interact with one another; and this interaction may lead to the inhibition of TGF-β signalling. Furthermore, the expression of these two genes, as described below, is specific to endothelial cells. Because these two genes 1) are both expressed specifically in endothelial cells, 2) are both up-regulated in endothelial cells under certain conditions, 3) encode MAD proteins that interact with one another in endothelial cells, and 4) inhibit TGF-β signalling (which is considered to be beneficial to atherosclerosis), rchd534 and fchd540 proteins are attractive targets for therapeutic intervention in cardiovascular disease. In particular, treatment regimens that inhibit the interaction or activity of the rchd534 and fchd540 proteins can be beneficial for the treatment cardiovascular disease.

Further analyses demonstrated that the rchd534 protein interacts with itself to form a homodimer. Thus, treatment regimens that inhibit the interaction of the rchd534 protein with itself can be beneficial for the treatment cardiovascular disease.

In addition, the analyses described below demonstrated novel interactions of both the rchd534 and fchd540 proteins with other proteins known to be involved in the TGF-β signalling pathway. The protein members of the TGF-β signalling pathway tested included MADR1 (Hoodless et al., 1996, Cell 85:489-500), MADR2 (Eppert et al., 1996, Cell 86: 543-552), DPC4 (Raftery et al., 1988, Genetics 139: 241-254), TβRI, TSR1, ActRIb, ALK3, and ALK6 (Wieser et al., 1995, EMBO J. 14: 2199-2208). For example, the rchd534 protein interacts strongly in endothelial cells with MADR1, MADR2, DPC4, and weakly in 293 (human embryonic kidney) cells with activated forms of receptors TβRI and ActRI. The fchd540 protein interacts strongly in 293 cells with activated forms of receptors TβRI and ALK6.

In the absence of transfected rchd543 and fchd540 genes, transfected MADR1 or transfected MADR2 mediated a 20-fold induction of a TGF-β inducible promoter in BAECs. Co-expression of either transfected rchd534 or transfected fchd540 in this system eliminated the induction, and also prevented the localization of MADR2 in the nucleus in response to TGF-β signalling. Therefore, treatment regimens that inhibit the interaction of the rchd534 and fchd540 proteins with other proteins involved in the TGF-β pathway also can be beneficial for the treatment cardiovascular of disease. As described above, the expression of rchd534 and fchd540 is specific, within arterial tissue, to endothelial cells. Accordingly, the rchd534 and fchd540 genes may be targets for intervention in a variety of inflammatory and fibroproliferative disorders that involve endothelial cells, including, but not limited to, cancer, angiogenesis, inflammation, and fibrosis.

15.1 MATERIALS AND METHODS

15.1.1. YEAST STRAINS, MEDIA, AND MICROBIOLOGICAL TECHNIQUES

Standard yeast media including synthetic complete medium lacking L-leucine, L-tryptophan, and L-histidine were prepared and yeast genetic manipulations were performed as described (Sherman, 1991, Meth. Enzymol., 194:3-21). Yeast transformations were performed using standard protocols (Gietz et al., 1992, Nucleic Acids Res., 20:1425. Ito et al., 1983, J. Bacteriol., 153:163-168). Plasmid DNAs were isolated from yeast strains by a standard method (Hoffman and Winston, 1987, Gene, 57:267-272).

15.1.2. PLASMID AND YEAST STRAIN CONSTRUCTION

The coding region of human fchd540 was amplified by PCR and cloned in frame into pGBT9 (Bartel et al., 1993, Cellular Interactions in Development. pp. 153-159) resulting in plasmid pGBT9-fchd540. pGBT9-fchd540 was transformed into two-hybrid screening strain HF7c and one resulting transformant was designated TB35.

15.1.3. TWO-HYBRID SCREENING

Two-hybrid screening was carried out essentially as described (Bartel et al., 1993, *supra*) using TB35 as the recipient strain and a human breast two-hybrid library.

15.1.4 PAPER FILTER BETA-GALACTOSIDASE ASSAYS

The paper filter beta-galactosidase (beta-gal) assay was performed essentially as previously described (Brill et al., 1994, Mol. Biol. Cell 5: 297-312).

15.2 RESULTS

15.2.1 STRONG PHYSICAL INTERACTION OF RCHD534 AND FCHD540 MEASURED BY TWO-HYBRID ASSAY

The fchd540 coding sequence was amplified by PCR and cloned into pGBT9 creating a GAL4 DNA-binding domain-fchd540 fusion gene. The screening strain HF7c was transformed with this construct. The rchd534 coding sequence was cloned into pGAD424 (Bartel et al., 1993, *supra*) creating a GAL4 transcriptional activation domain-rchd534 fusion gene, which was then used to transform strain Y187.

Yeast expression plasmids encoding the GAL4 DNA-binding domain either alone or fused in frame to fchd540, rchd534, *Drosophilia* MAD, DPC4, or p53 were transformed into MATa two-hybrid screening strain HF7c. Yeast expression plasmids encoding the GAL4 transcriptional activation domain alone and GAL4 activation domain fusions to rchd534 and SV40 were transformed into MATα two-hybrid screening strain Y187. p53 and SV40 interact with each other and should not interact with the experimental proteins. The HF7c transformants were propagated as stripes on semisolid synthetic complete medium lacking L-tryptophan and the Y187 transformants were grown as stripes on semisolid synthetic complete medium lacking L-leucine. Both sets of stripes were replica plated in the form of a grid onto a single rich YPAD plate and the haploid strains of opposite mating types were allowed to mate overnight at 30°C. The yeast strains on the mating plate were then replica plated to a synthetic complete plate

lacking L-leucine and L-tryptophan to select for diploids and incubated at 30°C overnight. Diploid strains on the synthetic complete plate lacking L-leucine and L-tryptophan were replica plated to a synthetic complete plate lacking L-leucine, L-tryptophan, and L-histidine to assay HIS3 expression and a paper filter on a synthetic complete plate lacking L-leucine and L-tryptophan. The next day the paper filter was subjected to the paper filter beta-galactosidase assay to measure expression of the lacZ reporter gene. HIS3 expression was scored after 3 days of growth at 30°C. The results are shown in Table 3.

The rchd534 fish protein was found to interact strongly with the fchd540 bait protein and not to interact with the rchd534, MAD, DPC4, p53, and GAL4 DNA binding domain bait proteins. This result demonstrated that rchd534 and fchd540 strongly physically interact with each other with significant specificity.

15.2.2 IDENTIFICATION OF PROTEINS THAT PHYSICALLY INTERACT WITH FCHD540

The fchd540 coding sequence was amplified by PCR and cloned into pGBT9 (Bartel et al., 1993, *supra*) creating a GAL4 DNA-binding domain-fchd540 fusion gene. HF7c was transformed with this construct resulting in strain TB35. TB35 grew on synthetic complete medium lacking L-tryptophan but not on synthetic complete medium lacking L-tryptophan and L-histidine demonstrating that the GAL4 DNA-binding domain-fchd540 fusion does not have intrinsic transcriptional activation activity.

TB35 was transformed with the human breast two-hybrid library and 5 million transformants were obtained. The transformants were plated on synthetic complete medium lacking L-leucine, L-tryptophan, and L-histidine and yeast colonies that both grew on synthetic complete medium lacking L-leucine, L-tryptophan, and L-histidine and expressed the beta-galactosidase reporter gene were identified. The 30 strains with the strongest beta-galactosidase induction were characterized. Library plasmids were isolated from these

strains, and the 5′ ends of all of the cDNA inserts were sequenced.

15.2.3 RETRANSFORMATION AND SPECIFICITY TESTING OF TCHV03A AND TCHVR4A

Two of the plasmids that encoded the strongest interactors were found to contain rchd534 cDNAs. Plasmid tchv03A was found to encode amino acids 17-235 of rchd534 and plasmid tchvR4A was found to encode amino acids 25-235 of rchd534.

It was confirmed that these rchd534 cDNAs encode proteins that physically interact specifically with fchd540. Yeast expression plasmids encoding the GAL4 DNA-binding domain either alone or fused in frame to fchd540, rchd534, Drosophila MAD, DPC4, and p53 were transformed into MATa two-hybrid screening strain HF7c. Yeast expression plasmids encoding the GAL4 transcriptional activation domain (GAL4 AD) alone and GAL4 activation domain fusions to tchv03a, tchvR4A and SV40 were transformed into MATα two-hybrid screening strain Y187. p53 and SV40 interact with each other and should not interact with the experimental proteins. The HF7c transformants were propagated as stripes on semi-solid synthetic complete medium lacking L-leucine. Both sets of stripes were replica plated in the form of a grid onto a single rich YPAD plate and the haploid strains of opposite mating types were allowed to mate overnight at 30°C. The yeast strains on the mating plate were then replica plated to a synthetic complete plate lacking L-leucine and L-tryptophan to select for diploids and incubated at 30°C overnight. Diploid strains on the synthetic complete plate lacking L-leucine and L-tryptophan were replica plated to a synthetic complete plate lacking L-leucine, L-tryptophan, and L-histidine to assay HIS3 expression and a paper filter on a synthetic complete plate lacking L-leucine and L-tryptophan. The next day the paper filter was subjected to the paper filter beta-galactosidase assay to measure expression of the lacZ reporter gene. HIS3 expression was scored after 3 days of growth at 30°C. The results are shown in the table below.

The strength or absence of physical interaction between each combination of test proteins is listed. Strong interactions are defined as interactions that cause the activation of both the HIS3 and lacZ reporter genes.

TABLE 3

	cDNA-GAL4 Activation Domain Fusion Tested				
	rchd534	tchv03A	tchvR4A	SV40	GAL4 AD alone
GAL4 DNA-Binding Domain Fusions					
fchd540	Strong	Strong	Strong	None	None
rchd534	None	None	None	None	None
Dros. MAD	None	None	None	None	None
DPC4	None	None	None	None	None
p53	None	None	None	Strong	None
GAL4 DNA-Binding Domain alone	None	None	None	None	None

The tchv03A and tchvR4A fish proteins were found to interact strongly with the fchd540 bait protein and to not interact with the rchd534, MAD, DPC4, p53, and GAL4 DNA binding domain bait proteins. These results confirm the result that the rchd534 and fchd540 proteins interact strongly with each other.

15.3 FURTHER ANALYSIS OF RCHD534 AND FCHD540 FUNCTION

The significance of the rchd534/fchd540 protein interaction was confirmed by examination of their expression and activity in human cells and animal models.

15.3.1 CHROMOSOMAL LOCALIZATION

The rchd534 gene was localized to chromosome 15 and the fchd540 gene was localized to chromosome 18, regions of the human genome that contain other MAD homologues. These regions of the human genome have also been implicated in the pathogenesis of several human malignancies.

15.3.2 TISSUE EXPRESSION PATTERNS

The expression patterns were examined using in situ hybridization techniques. Fluorescently labeled DNA probes of both the rchd534 and fchd540 genes were used to probe human carotid endartectomy samples. The expression of rchd534 and fchd540 was specific to endothelial cells lining the luminal surface of the carotid artery. In addition, a rabbit polyclonal antiserum generated against the rchd534 gene product prominently and selectively stained the endothelium present in large vessels such as human coronary arteries as well as smaller vessels present within human myocardium. Neither gene showed expression in any other cell type present in the arterial tissue sample, including smooth muscle cells and macrophages.

Expression patterns of both genes were also examined in response to certain stimulus. Both genes are selectively upregulated under the steady laminar shear stress (LSS) paradigm, but not under the turbulent shear stress paradigm or in response to stimulus by the cytokines rhIL-1β, TNFα, IFNγ or active TGFβ as measured in HUVEC cells. Thus, the rchd534 and the fchd540 genes appear to be selectively responsive to a LSS stimulus, manifesting no response to a non-laminar fluid mechanical stimulus, nor any other humoral stimuli tested. Thus, given that these two genes are: (1) localized to a region of the human genome that has been implicated in the pathogenesis of several human malignancies; (2) specifically expressed in a cell-type that is found only in vascular tissue, including atherosclerotic plaques; (3) up-regulated under the steady laminar shear stress cardiovascular disease paradigm; and (4) specifically inhibit TGF-β signalling indicate that rchd534 and fchd540 are excellent and specific targets for therapeutic intervention in the treatment of fibroproliferative and oncogenic disorders including tumor growth and vascularization.

15.3.3. CELLULAR LOCALIZATION

The cellular localization of the rchd534 and fchd540 proteins in bovine aortic endothelial cells (BAECs) was examined in relationship to other proteins involved in the TGF-β signalling pathway. In all experiments, the rchd534 and fchd540 proteins were located in the cytoplasm. MADR2 was located in the cytoplasm when transfected alone and in the nucleus when co-transfected with activated TβRI or when TGF-β was added to the culture medium. Co-transfection of rchd534 or fchd540 with MADR2 prevented the localization of MADR2 in the nucleus in response to TGF-β signalling.

15.3.4. PROTEIN INTERACTIONS IN HUMAN CELLS

The interaction of the rchd534 and fchd540 proteins, observed in yeast cells as described above, was tested in mammalian endothelial cell tissue culture. Either bovine aortic endothelial cells (BAECs) or 293 cells (human embryonic kidney cells, ATCC Accession No. CRL-1573) were transfected with constructs encoding both the rchd534 and fchd540 proteins, each fused to a different flag peptide allowing for specific immunoprecipitation. The rchd534 and fchd540 proteins were found to co-immunoprecipitate as heterodimers in extracts produced from both 293 cells and BAECs. The co-immunoprecipitation of rchd534 and fchd540 further supports that these proteins interact in human cells that are physiologically relevant to cardiovascular disease.

The ability of the rchd534 and fchd540 proteins to interact with themselves and with other protein members of the TGF-β signalling pathway (MADR1, MADR2, DPC4, TbR1, TSR1, ActR1b, ALK3, ALK6), was tested using this co-immunoprecipitation method. Each gene was transfected alone and in various combinations with other TGF-β pathway genes in either 293 cells or BAECs. The rchd534 protein formed homodimers in 293 cells and BAECs. The fchd540 protein did not form homodimers in 293 cells or BAECs. As mentioned above, the rchd534 and fchd540 proteins formed heterodimers in 293 cells and BAECs. This interaction is about 50 fold

stronger in BAECs than 293 cells based on equal amounts of protein. However, the rchd534-fchd540 protein interaction was significantly less avid than the rchd534 protein's interaction with itself.

The rchd534 protein interacted with MADR1, MADR2, and DPC4 in 293 cells and BAECs. The strength of MADR1 and MADR2 interactions was about the same between 293 cells and BAECs and much greater in BAECs for DPC4. The fchd540 protein interacted very weakly with MADR1, MADR2, and DPC4 in 293 cells. The rchd534 protein interacted strongly with activated forms of TβRI and ActRI and weakly with activated ALK6 in 293 cells. The fchd540 protein interacted strongly with activated TβRI and ALK6 receptors, and weakly with activated forms of TSRI, ALK3, and ActRIb in 293 cells. Thus, in addition to the interaction of the rchd534 and fchd540 proteins, the interaction of the rchd534 protein with itself, as well as the interaction of the rchd534 protein and the fchd540 protein with the other proteins in the TGF-β pathway described above are excellent targets for therapeutic intervention.

15.3.5 EFFECT OF EXPRESSION ON TGF-B SIGNALLING

The effect of both rchd534 and fchd540 on the TGF-β signalling pathway was tested in vitro. Primary BAECs were transfected with a construct called p3TP-Lux, containing a TGF-β responsive promoter fused to a reporter gene (Wrana et al., 1994, Nature 370: 341-347). The rchd534 gene or the fchd540 gene in pCI expression vectors (Promega) was transfected with and without MADR1 (pCMV5MADR1-Flag, Hoodless et al. 1996 Cell 85: 489-500) or MADR2 (pCMV5MADR2-Flag, Eppert et al. 1996 Cell 86: 543-552). The TGF-β response was induced 20-fold by either MADR1 or MADR2. Co-expression of either rchd534 or fchd540 completely eliminated this induction. Thus, the rchd534 and fchd540 proteins inhibited MADR1- and MADR2-mediated TGF-β signalling in endothelial cells. To confirm the specificity of this inhibitory effect, site specific mutants of both rchd534 or fchd540 were

constructed, based on known mutations identified in Drosophila homologues, that would be predicted to disrupt MAD-like signaling functions (Sekelsky et al., 1995, Genetics 139:1347-58; Raftery, 1995, Genetics 139:241-54; Newfeld et al., 1996, Development 122:2099-108; Wiersdorff et al., 1996, Development 122:2153-62). Unlike wild type rchd534 and fchd540, these mutant proteins were unable to inhibit the activation of the p3TP promoter in response to TGF-β. The expression levels of the mutant and wild-type proteins were comparable indicating the loss of function was not due to secondary instability.

Interestingly, Smad3, the *C. elegans* homolog to MAD3 which also functions in TGFβ signalling is over 90% identical to Smad2, the *C. elegans* MAD2 homolog, in the MH2 domain. Although this has not yet been directly investigated, it is likely that Smad7, the *C. elegans* homolog of the fchd540 gene, may function similarly to its inhibition to prevent association and activation of Smad3 by the TGFβ receptor, that is, to inhibit the phosphorylation of Smad3 and its association with protein components of the TGF-β signalling pathway.

These results further demonstrate that the interactions of either the rchd534 protein or the fchd540 protein with MADR2 or with activated TβR1 are excellent targets for therapeutic intervention. As described above, the expression of rchd534 and fchd540 is specific, within arterial tissue, to endothelial cells. Accordingly, the rchd534 and fchd540 genes may be targets for intervention in a variety of inflammatory and fibroproliferative disorders that involve endothelial cells, including, but not limited to, cancer angiogenesis, inflammation, and fibrosis.

16. EXAMPLE: THE RCHD534-LONG PROTEIN

As described below, the rchd534 gene was discovered to encode two spliceoforms. The short spliceoform, and the rchd534 protein it encodes, are described in detail in Section 9, above. A second novel spliceoform, encoding a novel longer protein designated rchd534-long, is described in detail in the subsections below.

16.1 IDENTIFICATION AND CHARACTERIZATION OF THE RCHD534-LONG SPLICEOFORM AND PROTEIN

A human heart cDNA library (Stratagene, LaJolla, CA) was screened with a probe containing nucleotides 400-700 of the fchd540 (see FIG.10) under the following hybridization conditions: hybridization overnight at 65°C, washing with 2xSSC and 0.1% SDS for 20 minutes at room temperature, followed two washes with 0.2xSSC and 0.1%SDS for 20 minutes at 65°C. A positive clone was found to encode a novel protein, related to the rchd534 protein, that was designated the rchd534-long protein. The rchd534-long protein, like the rchd534 protein, has an MH2 domain. In addition, the rchd534-long protein has an MH1 domain not present in rchd534. The original clone isolated from the Stratagene heart library contained incorrect sequence, including two stop codons, between the MH1 and MH2 domain coding regions. Therefore, cDNA prepared from human heart mRNA (Clontech, Palo Alto, CA) was used as template for the PCR reaction to isolate the correct cDNA sequence in the region spanning the MH1 and MH2 domain coding regions. The following two pairs of nested primers were used:

Pair 1: A) 5-'GAGGCTGCGGCCGCTCCGAAGTCC-3'
B) 5'-CTCCGCCGGGGCCGCCACTATCT-3'

Pair 2: A) 5'-CCGGGACGCAGTGGGACAG-3'
B) 5'-CGGGGAGTTGACGAAGATGG-3'

The nucleotide sequence of the PCR amplification product was determined and confirmed to be accurate by sequence analysis of several amplification products from both

Clontech human lung cDNA, and heart mRNA that was obtained from Clontech and then reverse-transcribed into cDNA. The incorrect sequence in the original clone obtained from the Stratagene heart library was replaced by a PCR amplification product containing the correct sequence.

This correct cDNA encoding the rchd534-long protein was cloned into the TA cloning vector (Invitrogen) to create plasmid pHL6TA1A, which was deposited with the American Type Culture Collection on February 6, 1998 as Accession No. 209615. The cDNA sequence of the rchd534-long spliceoform encoding the entire rchd534-long protein is shown in FIG.9. The rchd534-long nucleotide sequence is 93% identical to the nucleotide sequence of the mouse SMAD6 gene (Imamura et al., 1997, Nature 389: 622-626). The rchd534-long sequence was reported in Hata et al., 1998, Genes and Development 12: 186-197.

The domains of the rchd534-long protein are shown in schematic form and compared with the domains of the rchd534 protein in FIG.11. The rchd534-long protein contains an MH1 domain, a spacer region, and an MH2 domain. The rchd534-long protein contains an N-terminal 273 amino acids (from Met-1 to Glu-273) which are not present in the rchd534 protein. This N-terminal region contains an MH1 domain that is more highly homologous to the MH1 domain of fchd540 than to the MH1 domain of other MAD proteins.

The amino acid sequence of rchd534-long protein from amino acid Ser-274 to the C-terminal amino acid Arg-496 are identical to amino acids Ser-13 to the C-terminal amino acid Arg-235 of the rchd534 protein. This region comprises an MH2 domain, from Pro-328 to Arg-496 in rchd534-long, and from Pro-67 to Arg-235 in rchd534. The rchd534 protein contains a 12 amino acid sequence at the N-terminus, from Met-1 to Lys-12, which is not present in the rchd534-long protein.

16.2. TGF-β SIGNALLING INHIBITORY ACTIVITY OF THE RCHD534-LONG PROTEIN

The activity of the rchd534-long protein was tested using the TGF-β responsive reporter system described in Section 15.3.5, above for the rchd534 and fchd540 proteins.

HEPG2, 293 and BAEC cells were transfected with a construct called p3TP-Lux, containing a TGF-β responsive promoter fused to a reporter gene (Wrana et al., 1994, Nature 370: 341-347). The rchd534-long gene in pCI expression vectors (Promega) was transfected with and without MADR2. Luciferase activity produced from the reporter construct was measured relative to an internal control (secreted alkaline phosphatase expression). The results obtained in 293 (human embryonic kidney) cells are summarized in Table 4 below. A "+" indicates that the specified construct, plasmid, or gene was present in the cell line assayed; whereas a "-" indicates that it was absent.

TABLE 4

3TP-LUX	TGF-bRI	pCI	MADR2	fchd540	rchd534	rchd534-long	Relative Luciferase Activity
+	+	-	-	-	-	-	10
+	+	-	-	-	-	-	40
+	+	-	-	-	-	-	40
+	+	-	-	-	-	-	107
+	+	-	-	+	-	-	2
+	+	-	-	-	+	-	10
+	+	-	-	-	-	+	2
+	+	+	+	-	-	-	104
+	+	-	+	-	-	-	124
+	+	-	+	+	-	-	13
+	+	-	+	-	+	-	59
+	+	-	+	-	-	+	18

The TGF-β response, both un-induced by MADR2 and induced by MADR2 overexpression, was inhibited by expression of rchd534-long, as well as by rchd534 and fchd540. Similar

inhibition of TGF-β signalling by the rchd534-long protein was observed in each of the several different cell lines.

17. EXAMPLE: ANTISENSE AND RIBOZYME MOLECULES FOR INHIBITION OF RCHD534 AND FCHD540 EXPRESSION

The principles presented in Section 5.6.1.1, above, can be used to design oligonucleotides for use in inhibiting the expression of target genes, such as the rchd534 or fchd540 genes.

17.1. ANTISENSE MOLECULES

The following antisense molecules can be used to inhibit translation of the rchd534 protein:

a) 5′-CATTTCATTTCATACAA-3′ which is complementary to nucleotides -14 to +3 of in FIG.8.

b) 5′-CATTTCATTTCATACAATATATG-3′ which is complementary to nucleotides -20 to +3 in FIG.8.

c) 5′-CATTTCATTTCATACAATATATGGCCTTT-3′ which is complementary to nucleotides -26 to +3 in FIG.8.

d) 5′-CATTTCATTTCATACAATATATGGCCTTTTGTGGC-3′ which is complementary to nucleotides -32 to +3 in FIG.8.

e) 5′-GGACATTTCATTTCATACAATATATGGCCTTTTGT-3′ which is complementary to nucleotides -29 to +6 in FIG.8.

f) 5′-TTCATTTCATACAATATATGGCCTTTTGT-3′ which is complementary to nucleotides -29 to -1 in FIG.8.

g) 5′-TCATACAATATATGGCCTTTTGT-3′ which is complementary to nucleotides -29 to -7 in FIG.8.

h) 5′-AATATATGGCCTTTTGT-3′ which is complementary to nucleotides -29 to -13 in FIG.8.

The following antisense molecules can be used to inhibit translation of the rchd534-long protein:

a) 5'-CATACGATATCCTTTGGCGCCAGGGG-3', which is complementary to nucleotides -23 TO +3 in FIG.9.

b) 5'-GGACCTGAACATACGATATCCTTTGGCGCCAGGGG-3', which is complementary to nucleotides -23 TO +12 in FIG.9.

c) 5'-CATACGATATCCTTTGGCGCCAGGGGTGGGGGGG-3', which is complementary to nucleotides -31 TO +3 in FIG.9.

The following antisense molecules can be used to inhibit translation of the fchd540 protein:

a) 5'-CATGCGGGGCGAGGAGG-3' which is complementary to nucleotides -14 to +3 of fchd540 in FIG.10.

b) 5'-CATGCGGGGCGAGGAGGCGAGGA-3' which is complementary to nucleotides -20 to +3 of fchd540 in FIG.10.

c) 5'-CATGCGGGGCGAGGAGGCGAGGAGAAAAG-3' which is complementary to nucleotides -26 to +3 of fchd540 in FIG.10.

d) 5'-CATGCGGGGCGAGGAGGCGAGGAGAAAAGTCGTTT-3' which is complementary to nucleotides -32 to +3 of fchd540 in FIG.10.

e) 5'-GAACATGCGGGGCGAGGAGGCGAGGAGAAAAGTCG-3' which is complementary to nucleotides -29 to +6 of fchd540 in FIG.10.

f) 5'-GCGGGGCGAGGAGGCGAGGAGAAAAGTCG-3' which is complementary to nucleotides -29 to -1 of fchd540 in FIG.10.

g) 5′-CGAGGAGGCGAGGAGAAAAGTCG-3′ which is complementary to nucleotides -29 to -7 of fchd540 in FIG.10.

h) 5′-GGCGAGGAGAAAAGTCG-3′ which is complementary to nucleotides -29 to -13 of fchd540 in FIG.10.

17.2. RIBOZYME MOLECULES

The central, catalytic portion of a hammerhead ribozyme molecule consist of the following sequence:

5′-*C*AAAGCNGNXXXXNCNGAGNAGUC-3′;

wherein the 5′-proximal CA bases hybridize to a complementary 5′-UG-3′ in the target mRNA. The first four underlined bases form a stem by base pairing with the second set of underlined bases, with the intervening bases, shown as X's, forming a non-pairing loop. In order to hybridize to a target mRNA, a hammerhead ribozyme contains additional bases flanking each end of the central segment shown above. The 5′ ribozyme flanking segment is complementary to the respective flanking sequences immediately 3′ to the target UG; and the 3′ flanking segment is complementary to the respective flanking sequence beginning two bases upstream of the target U, and extending 5′-ward (in effect, skipping the first base upstream of the target U). Cleavage occurs between first and second bases upstream of (*i*.e., 5′ to) the U in the target 5′-UG-3′ site.

The following ribozyme molecules can be used to inhibit translation of the rchd534 protein:

a) 5′-GGUGGAGCCCCAGGGCAUUACCUCAAAGCNGNXXXXNCNGAGNAGUCGUGG GCAAGGUGGGCACUCAGGUGGG-3′ which will cleave the short spliceoform rchd534 mRNA between nucleotides 716 and 717 in FIG.8.

b) 5′-GUGUCUCUAUGGGUUUGCCCAAAGCNGNXXXXNCNGAGNAGUCUCUGGACA UUUCAUUUCAUAC-3′ which will cleave the short

spliceoform rchd534 mRNA between nucleotides 1040 and 1041 in FIG.8.

The following ribozyme molecules can be used to inhibit translation of the rchd534 protein or the rchd534-long protein:

a) 5'-GGCCCUCUCGCCGUCGGGCUCCUUGCUGAGCAAAGCNGNXXXXNCNGAGNA GUCGAUGCCGAAGCCGAUCUUGCUGCGCG-3' which will cleave the short spliceoform rchd534 mRNA between nucleotides 1421 and 1422 in FIG.8; and the rchd534-long spliceoform mRNA and between nucleotides 1327 and 1328 in FIG.9.

The following ribozyme molecules can be used to inhibit translation of the rchd534-long protein:

a) 5'-CGACUUCGCCAAAGUCGCCGCAAAGCNGNXXXXNCNGAGNAGUCCAGCCCC GAGCGTTTGGACCTG-3', which will cleave between nucleotides +178 and +179 in FIG. 9.

The following ribozyme molecules can be used to inhibit translation of the fchd540 protein:

a) 5'-CGUUUGCCUGCUAAGGAGCGAACAAAGCNGNXXXXNCNGAGNAGUCGAUGU UUCUUUUGUGAGUCGGGCGCCG-3', which will cleave the fchd540 mRNA between nucleotides -53 and -52 in FIG.10.

b) 5'-CGCCGGACGAGCGCAGAUCGUUUGGUCCUGAACAAAGCNGNXXXXNCNGAG NAGUCCGGGGCGAGGAGGCGAGGAGAAAAGUCG-3', which will cleave the fchd540 mRNA between nucleotides -1 and +1 in FIG.10.

c) 5'-GGAGUAAGGAGGGGGGGGAGACUCUAGUUCGCAAAGCNGNXXXXNCNGAGN AGUCAGUCGGCUAAGGUGAUGGGGGUUGCAGCACACC-3' which will cleave the fchd540 mRNA between nucleotides +602 and +603 in FIG.10.

18. DEPOSIT OF MICROORGANISMS

The following microorganisms were deposited with the Agricultural Research Service Culture Collection (NRRL), Peoria, Illinois, on January 11, 1995 and assigned the indicated accession numbers:

Microorganism	NRRL Accession No.
RCHD005	B-21376
RCHD024	B-21377
RCHD032	B-21378
RCHD036	B-21379
RCHD502	B-21380
RCHD523	B-21381
RCHD528	B-21382

The following microorganisms were deposited with the Agricultural Research Service Culture Collection (NRRL), Peoria, Illinois, on June 6, 1995 and assigned the indicated accession numbers:

Microorganism	NRRL Accession No.
FCHD523	B-21458
FCHD534	B-21459

The following microorganisms were deposited with the American Type Culture Collection (ATCC), Rockville, Maryland, on February 7, 1996, and assigned the indicated accession numbers:

Microorganism	ATCC Accession No.
FCHD502SF	69981
FCHD502SJ	69982

The following microorganisms were deposited with the American Type Culture Collection (ATCC), Rockville, Maryland, on February 9, 1996, and assigned the indicated accession numbers:

Microorganism	ATCC Accession No.
FCHD528A	69985
FCHD528B	69986
FCHD528C	69987

The following microorganism was deposited with the American Type Culture Collection (ATCC), Rockville, Maryland, on February 7, 1996, and assigned the indicated accession number:

Microorganism	ATCC Accession No.
pFCHD540	69984

The following plasmid, encoding the rchd534-long protein, was deposited with the American Type Culture Collection (ATCC), Rockville, Maryland, on February 6, 1998, and assigned the indicated accession number:

Microorganism	ATCC Accession No.
pHL6TA1A	209615

The present invention is not to be limited in scope by the specific embodiments described herein, which are intended as single illustrations of individual aspects of the invention, and functionally equivalent methods and components are within the scope of the invention. Indeed, various modifications of the invention, in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and accompanying drawings. Such modifications are intended to fall within the scope of the appended claims.

WHAT IS CLAIMED IS:

1. An isolated polynucleotide comprising a nucleotide sequence (a) encoding a polypeptide having the amino acid sequence set forth in FIG.9, or (b) encoding a polypeptide encoded by the rchd534-long cDNA contained in the clone pHL6TA1A, as deposited with the American Type Culture Collection as Accession No. 209615, or (c) which is the complement of (a) or (b).

2. An isolated polynucleotide comprising the nucleotide sequence (a) of the rchd534-long cDNA as shown in FIG.9, or (b) of the rchd534 cDNA insert contained in the clone pHL6TA1A, as deposited with the American Type Culture Collection as Accession No. 209615, or (c) which is the complement of (a).

3. An isolated polynucleotide that hybridizes under highly stringent conditions to the nucleotide sequence of Claim 1.

4. An isolated polynucleotide that encodes a protein member of the TGF-β signalling pathway, wherein the polynucleotide hybridizes under moderately stringent conditions to the nucleotide sequence of Claim 1.

5. An isolated polynucleotide comprising the nucleotide sequence (a) of the rchd534-long polypeptide coding region, which coding region is set forth from nucleotide residue number 155 to 494 of FIG.9, or (b) of the polypeptide coding region of the rchd534-long cDNA contained in the clone pHL6TA1A, as deposited with the American Type Culture Collection as Accession No. 209615, or (c) which is the complement of (a) or (b).

6. An isolated polynucleotide that hybridizes under highly stringent conditions to the nucleotide sequence of Claim 5.

7. An isolated polynucleotide that encodes a protein member of the TGF-β signalling pathway, wherein the polynucleotide hybridizes under moderately stringent conditions to the nucleotide sequence of Claim 5.

8. The isolated polynucleotide of Claims 1, 2, 3, 4, 5, 6, or 7 which is DNA.

9. The isolated polynucleotide of Claim 8 which is cDNA.

10. The isolated polynucleotide of Claim 8 which is genomic DNA.

11. The isolated polynucleotide of Claims 1, 2, 3, 4, 5, 6, or 7 which is RNA.

12. The isolated polynucleotide of Claims 1, 2, 3, 4, 5, 6, or 7 which further comprises a detectable label.

13. A vector containing the polynucleotide of Claims 1, 2, 3, 4, 5, 6, or 7.

14. An expression vector containing the polynucleotide of Claims 1, 2, 3, 4, 5, 6, or 7 in operative association with a nucleotide regulatory element that controls expression of the polynucleotide in a host cell.

15. A cultured genetically engineered host cell containing the polynucleotide of Claims 1, 2, 3, 4, 5, 6, or 7.

16. A cultured genetically engineered host cell containing the polynucleotide of Claims 1, 2, 3, 4, 5, 6, or 7 in operative association with a nucleotide regulatory element that controls expression of the polynucleotide in the host cell.

17. The genetically engineered host cell of Claim 16 which is prokaryotic.

18. The genetically engineered host cell of Claim 16 which is eukaryotic.

19. A method of producing an rchd534-long polypeptide, comprising the steps of:

(a) growing the genetically engineered host cell of Claim 17 in a culture; and

(b) collecting the polypeptide from the culture.

20. A method of producing an rchd534-long polypeptide, comprising the steps of:

(a) growing the genetically engineered host cell of Claim 18 in a culture; and

(b) collecting the polypeptide from the culture.

21. A method for identifying a substance for treating cardiovascular disease comprising assaying the ability of the substance to modulate the expression of the rchd534 gene, or the activity of the rchd534 or rchd534-long protein.

22. The method of Claim 21 in which the cardiovascular disease is atherosclerosis.

23. The method of Claim 21 in which the cardiovascular disease is ischemia/reperfusion.

24. The method of Claim 21 in which the cardiovascular disease is hypertension.

25. The method of Claim 21 in which the cardiovascular disease is restenosis.

26. The method of Claim 21 in which the modulation of the expression of said gene is assayed by:

(a) exposing a sample of cells to a test substance;

(b) assaying the expression of said gene in the sample of cells; and

(c) comparing the expression level of the gene in the sample exposed to the substance to the expression level of the gene in a control sample of cells, in which a difference between the expression level of the gene in the sample exposed to the substance and the control indicates the modulation of expression of the gene.

27. The method of Claim 26 in which the gene is down-regulated by the test substance.

28. The method of Claim 27 in which the substance is an oligonucleotide complementary to the 5′ region of the gene and blocks transcription via triple helix formation.

29. The method of Claim 27 in which the substance is an antisense or ribozyme molecule that blocks translation of the gene.

30. The method of Claim 26 in which the gene is up-regulated by the test substance.

31. The method of claim 21 in which the substance is a small organic or inorganic molecule that modulates the activity of the protein product by binding to the protein product.

32. The method of claim 21 in which the substance is an antibody that modulates the activity of the protein product by binding to the protein product.

33. An assay for identifying a substance that binds to the rchd534-long protein, comprising:

(a) contacting a protein or peptide containing an amino acid sequence corresponding to the binding site of the protein with a test substance, under conditions and for a time sufficient to permit binding and formation of a complex between the protein or peptide and the test substance, and

(b) detecting the formation of a complex, in which the ability of the test substance to bind to the protein is indicated by the presence of the test substance in the complex.

34. An assay for identifying a substance that inhibits the interaction between the rchd534-long protein and the fchd540 protein comprising:

(a) contacting a protein or peptide containing an amino acid sequence corresponding to the binding site of the rchd534-long protein with a protein or peptide containing an amino acid sequence corresponding to the binding site of the fchd540 protein, under conditions and for a time sufficient to permit binding and formation of a complex, in the presence of a test substance, and

(b) detecting the formation of a complex, in which the ability of a test substance to inhibit the interaction between the rchd534-long protein and fchd540 protein is indicated by a decrease in complex formation as compared to the amount of complex formed in the absence of the test substance.

35. An assay for identifying a substance that inhibits the interaction between two rchd534-long protein molecules comprising:

(a) contacting a first protein or peptide containing an amino acid sequence corresponding to the binding site of the rchd534-long protein with a second protein or peptide containing an amino acid sequence corresponding to the binding site of the rchd534-long protein, under conditions and for a time sufficient to permit binding and formation of a complex, in the presence of a test substance, and

(b) detecting the formation of a complex, in which the ability of a test substance to inhibit the interaction between two rchd534-long protein molecules is indicated by a decrease in complex formation as compared to the amount of complex formed in the absence of the test substance.

36. An assay for identifying a substance that inhibits the interaction between the rchd534-long protein and a protein member of the TGF-β signalling pathway comprising:

(a) contacting a protein or peptide containing an amino acid sequence corresponding to the binding site of the rchd534-long protein with a protein or peptide containing an amino acid sequence corresponding to the binding site of the protein member of the TGF-β signalling pathway, under conditions and for a time sufficient to permit binding and formation of a complex, in the presence of a test substance, and

(b) detecting the formation of a complex, in which the ability of a test substance to inhibit the interaction between the rchd534-long protein and the protein member of the TGF-β signalling pathway is indicated by a decrease in complex formation as compared to the amount of complex formed in the absence of the test substance.

37. The assay of Claim 36 wherein the protein member of the TGF-β signalling pathway is MADR1, MADR2, DPC4, activated TβR1, activated ActR1b, or activated ALK6.

38. An assay for identifying a substance that inhibits the interaction between the fchd540 protein and a protein member of the TGF-β signalling pathway comprising:

(a) contacting a protein or peptide containing an amino acid sequence corresponding to the binding site of the fchd540 protein with a protein or peptide containing an amino acid sequence corresponding to the binding site of the protein member of the TGF-β signalling pathway, under conditions and for a time sufficient to permit binding and

formation of a complex, in the presence of a test substance, and

(b) detecting the formation of a complex, in which the ability of the test substance to inhibit the interaction between the fchd540 protein and the protein member of the TGF-β signalling pathway is indicated by a decrease in complex formation as compared to the amount of complex formed in the absence of the test substance.

39. The assay of Claim 38 wherein the protein member of the TGF-β signalling pathway is MADR1, MADR2, DPC4, activated TβR1, activated ALK6, activated TSR1, activated ALK3, or activated ActR1β.

40. A method for treating cardiovascular disease comprising administering a compound that inhibits the interaction between the rchd534-long protein and the fchd540 protein.

41. A method for treating cardiovascular disease comprising administering a compound that inhibits the interaction between two rchd534-long protein molecules.

42. A method for treating cardiovascular disease comprising administering a compound that inhibits the interaction between the rchd534-long protein and a protein member of the TGF-β signalling pathway.

43. The method of Claim 42 wherein the protein member of the TGF-β signalling pathway is MADR1, MADR2, DPC4, activated TβR1, activated ActR1b, or activated ALK6.

44. A method for treating cardiovascular disease comprising administering a compound that inhibits the interaction between the fchd540 protein and a protein member of the TGF-β signalling pathway.

45. The method of Claim 44 wherein the protein member of the TGF-β signalling pathway is MADR1, MADR2, DPC4, activated TβR1, activated ALK6, activated TSR1, activated ALK3, or activated ActR1β.

46. A method for identifying a substance that enhances the TGF-β signalling response comprising:

(a) contacting a genetically engineered cell with a test substance, said cell comprising 1) a reporter gene in operative association with an inducible TGF-β regulatory element; 2) a recombinant gene encoding the rchd534-long protein or a recombinant gene encoding the fchd540 protein; and 3) a recombinant gene encoding the MADR1 protein or a recombinant gene encoding the MADR2 protein; and

(b) detecting expression of said reporter gene in which ability of the test substance to enhance the TGF-β signalling response is indicated by an increase in expression of the reporter gene as compared to the amount of expression in the absence of the test substance.

47. A method for identifying a substance for treating fibroproliferative disease or oncogenic related disorders comprising assaying the ability of the substance to modulate expression of, or the activity of the encoded protein product of, the rchd534-long spliceoform or the fchd540 gene.

48. The method of Claim 47 in which the fibroproliferative disease is diabetic retinopathy.

49. The method of Claim 47 in which the oncogenic related disorder is a tumor growth.

50. The method of Claim 47 in which the oncogenic related disorder is angiogenesis.

51. A method for treating fibroproliferative disease or oncogenic related disorders comprising administering a

compound that inhibits the interaction between the rchd534-long protein and a protein member of the TGF-β signalling pathway.

52. A method for treating fibroproliferative disease or compound that inhibits the interaction between the rchd534-long protein and the fchd540 protein.

ABSTRACT OF THE INVENTION

The present invention relates to methods and compositions for the treatment and diagnosis of cardiovascular disease, including, but not limited to, atherosclerosis, ischemia/reperfusion, hypertension, restenosis, and arterial inflammation. Specifically, the present invention identifies and describes genes which are differentially expressed in cardiovascular disease states, relative to their expression in normal, or non-cardiovascular disease states, and/or in response to manipulations relevant to cardiovascular disease. Further, the present invention identifies and describes genes via the ability of their gene products to interact with gene products involved in cardiovascular disease. Still further, the present invention provides methods for the identification and therapeutic use of compounds as treatments of cardiovascular disease. Moreover, the present invention provides methods for the diagnostic monitoring of patients undergoing clinical evaluation for the treatment of cardiovascular disease, and for monitoring the efficacy of compounds in clinical trials. Additionally, the present invention describes methods for the diagnostic evaluation and prognosis of various cardiovascular diseases, and for the identification of subjects exhibiting a predisposition to such conditions.

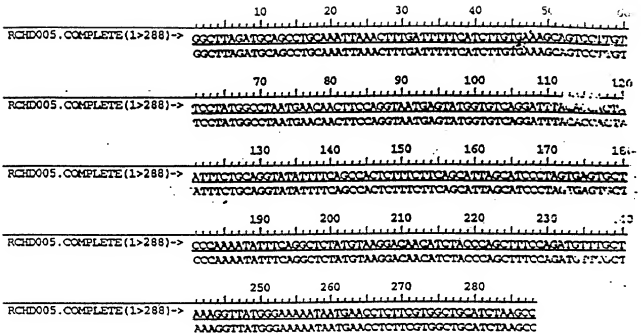
RCHD005.COMPLETE(1>288)->
GGCTTAGATGCAGCCTGCAAATTAAACTTTGATTTTTCATCTTGTGAAAGCAGTCCTAGT
GGCTTAGATGCAGCCTGCAAATTAAACTTTGATTTTTCATCTTGTGAAAGCAGTCCTAGT
RCHD005.COMPLETE(1>288)->
TCCTATGGCCTAATGAACAACTTCCAGGTAATGAGTATGGTGTCAGGATTTACACCAGTA
TCCTATGGCCTAATGAACAACTTCCAGGTAATGAGTATGGTGTCAGGATTTACACCAGTA
RCHD005.COMPLETE(1>288)->
ATTTCTGCAGGTATATTTTCAGCCACTCTTTCTTCAGCATTAGCATCGCTAGTGAGTGCT
ATTTCTGCAGGTATATTTTCAGCCACTCTTTCTTCAGCATTAGCATCGCTAGTGAGTGCT
RCHD005.COMPLETE(1>288)->
CCCAAAATATTTCAGGCTCTATGTAAGGACAACATCTACCCAGCTTTCCAGATCTTTGCT
CCCAAAATATTTCAGGCTCTATGTAAGGACAACATCTACCCAGCTTTCCAGATCTTTGCT
RCHD005.COMPLETE(1>288)->
AAAGGTTATGGGAAAAATAATGAACCTCTTCGTGGCTGCATCTAAGCC
AAAGGTTATGGGAAAAATAATGAACCTCTTCGTGGCTGCATCTAAGCC

FIG. 1

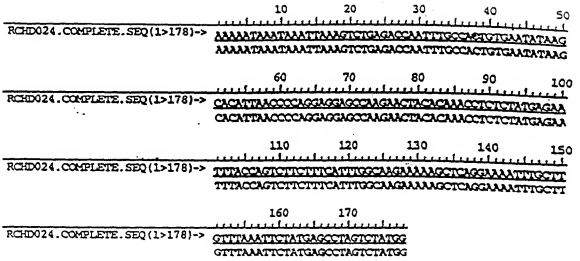
10 20 30 40 50
RCHD024.COMPLETE.SEQ(1>178)-> AAAAATAAATAAATTAAAGTCTGAGACCAATTTGCCACTGTGAATATAAG
AAAAATAAATAAATTAAAGTCTGAGACCAATTTGCCACTGTGAATATAAG
60 70 80 90 100
RCHD024.COMPLETE.SEQ(1>178)-> CACATTAACCCCAGGAGGAGCCAAGAACTACACAAACCTCTCTATGAGAA
CACATTAACCCCAGGAGGAGCCAAGAACTACACAAACCTCTCTATGAGAA
110 120 130 140 150
RCHD024.COMPLETE.SEQ(1>178)-> TTTACCAGTCTTCTTTCATTTGGCAAGAAAAAGCTCAGGAAAATTTGCTT
TTTACCAGTCTTCTTTCATTTGGCAAGAAAAAGCTCAGGAAAATTTGCTT
160 170
RCHD024.COMPLETE.SEQ(1>178)-> GTTTAAATTCTATGAGCCTAGTCTATGG
GTTTAAATTCTATGAGCCTAGTCTATGG

FIG. 2

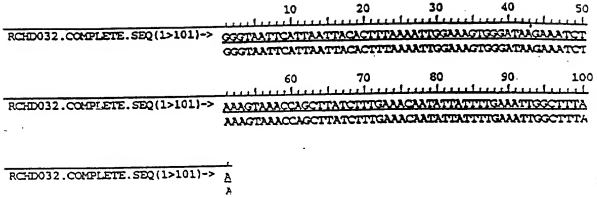
10 20 30 40 50
RCHD032.COMPLETE.SEQ(1>101)-> GGGTAATTCATTAATTACACTTTAAAATTGGAAAGTGGGATAAGAAATCT
GGGTAATTCATTAATTACACTTTAAAATTGGAAAGTGGGATAAGAAATCT
60 70 80 90 100
RCHD032.COMPLETE.SEQ(1>101)-> AAAGTAAACCAGCTTATCTTTGAAACAATATTATTTTGAAATTGGCTTTA
AAAGTAAACCAGCTTATCTTTGAAACAATATTATTTTGAAATTGGCTTTA
RCHD032.COMPLETE.SEQ(1>101)-> A
A

FIG. 3

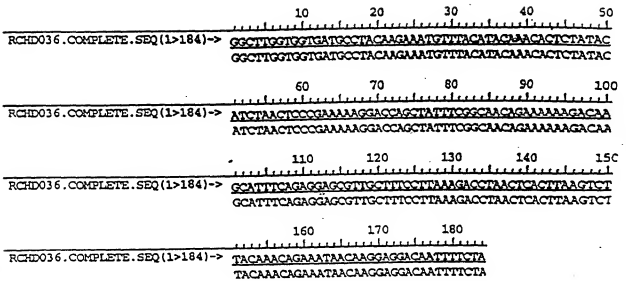
10 20 30 40 50
RCHD036.COMPLETE.SEQ(1>184)-> GGCTTGGTGGTGATGCCTACAAGAAATGTTTACATACAAACACTCTATAC
GGCTTGGTGGTGATGCCTACAAGAAATGTTTACATACAAACACTCTATAC
60 70 80 90 100
RCHD036.COMPLETE.SEQ(1>184)-> ATCTAACTCCCGAAAAAGGACCAGCTATTTCGGCAACAGAAAAAAGACAA
ATCTAACTCCCGAAAAAGGACCAGCTATTTCGGCAACAGAAAAAAGACAA
110 120 130 140 150
RCHD036.COMPLETE.SEQ(1>184)-> GCATTTCAGAGGAGCGTTGCTTTCCTTAAAGACCTAACTCACTTAAGTCT
GCATTTCAGAGGAGCGTTGCTTTCCTTAAAGACCTAACTCACTTAAGTCT
160 170 180
RCHD036.COMPLETE.SEQ(1>184)-> TACAAACAGAAATAACAAGGAGGACAATTTTCTA
TACAAACAGAAATAACAAGGAGGACAATTTTCTA

FIG. 4

```
M  G  L  L  P  K  L  G  A  S  Q  G  S  D  T  S  T  S  R  A      20
ATG GGG CTC CTG CCC AAG CTC GGC GCG TCC CAG GGC AGC GAC ACC TCT ACT AGC CGA GCC      60

G  R  C  A  R  S  V  F  G  N  I  K  V  F  V  L  C  Q  G  L      40
GGC CGC TGT GCC CGC TCG GTC TTC GGC AAC ATT AAG GTG TTT GTG CTC TGC CAA GGC CTC     120

L  Q  L  C  Q  L  L  Y  S  A  Y  F  K  S  S  L  T  T  I  E      60
CTG CAG CTC TGC CAA CTC CTG TAC AGC GCC TAC TTC AAG AGC AGC CTC ACC ACC ATT GAG     180

K  R  F  G  L  S  S  S  S  S  G  L  I  S  S  L  N  E  I  S      80
AAG CGC TTT GGG CTC TCC AGT TCT TCA TCG GGT CTC ATT TCC AGC TTG AAT GAG ATC AGC     240

N  A  I  L  I  I  F  V  S  Y  F  G  S  R  V  H  R  P  R  L     100
AAT GCC ATC CTC ATC ATC TTT GTC AGC TAC TTT GGC AGC CGG GTG CAC CGT CCA CGT CTG     300

I  G  I  G  G  L  F  L  A  A  G  A  F  I  L  T  L  P  H  F     120
ATT GGC ATC GGA GGT CTC TTC CTG GCT GCA GGT GCC TTC ATC CTC ACC CTC CCA CAC TTC     360

L  S  E  P  Y  Q  Y  T  L  A  S  T  G  N  N  S  R  L  Q  A     140
CTC TCC GAG CCC TAC CAG TAC ACC TTG GCC AGC ACT GGG AAC AAC AGC CGC TTG CAG GCC     420

E  L  C  Q  K  H  W  Q  D  L  P  P  S  K  C  H  S  T  T  Q     160
GAG CTC TGC CAG AAG CAT TGG CAG GAC CTG CCT CCC AGT AAG TGC CAC AGC ACC ACC CAG     480

N  P  Q  K  E  T  S  S  M  W  G  L  M  V  V  A  Q  L  L  A     180
AAC CCC CAG AAG GAG ACC AGC AGC ATG TGG GGC CTG ATG GTG GTT GCC CAG CTG CTG GCT     540

G  I  G  T  V  P  I  Q  P  F  G  I  S  Y  V  D  D  F  S  E     200
GGC ATC GGG ACA GTG CCT ATT CAG CCA TTT GGG ATC TCC TAT GTG GAT GAC TTC TCA GAG     600

P  S  N  S  P  L  Y  I  S  I  L  F  A  I  S  V  F  G  P  A     220
CCC AGC AAC TCG CCC CTG TAC ATC TCC ATC TTA TTT GCC ATC TCT GTA TTT GGA CCG GCT     660

F  G  Y  L  L  G  S  V  M  L  Q  I  F  V  D  Y  G  R  V  N     240
TTC GGG TAC CTG CTG GGC TCT GTC ATG CTG CAG ATC TTT GTG GAC TAT GGC AGG GTC AAC     720

T  A  A  V  N  L  V  P  G  D  P  R  W  I  G  A  W  W  L  G     260
ACA GCT GCA GTT AAC TTG GTC CCG GGT GAC CCC CGA TGG ATT GGA GCC TGG TGG CTA GGC     780

L  L  I  S  S  A  L  L  V  L  T  S  F  P  F  F  F  F  P  R     280
CTG CTC ATT TCT TCA GCT TTA TTG GTT CTC ACC TCT TTC CCC TTT TTT TTC TTC CCT CGA     840

A  M  P  I  G  A  K  R  A  P  A  T  A  D  E  A  R  K  L  E     300
GCA ATG CCC ATA GGA GCA AAG AGG GCT CCT GCC ACA GCA GAT GAA GCA AGG AAG TTG GAG     900

E  A  K  S  R  G  S  L  V  D  F  I  K  R  F  P  C  I  F  L     320
GAG GCC AAG TCA AGA GGC TCC CTG GTG GAT TTC ATT AAA CGG TTT CCA TGC ATC TTT CTG     960

R  L  L  M  N  S  L  F  V  L  V  V  L  A  Q  C  T  F  S  S     340
AGG CTC CTG ATG AAC TCA CTC TTC GTC CTG GTG GTC CTG GCC CAG TGC ACC TTC TCC TCC    1020

V  I  A  G  L  S  T  F  L  N  K  F  L  E  K  Q  Y  G  T  S     360
GTC ATT GCT GGC CTC TCC ACC TTC CTC AAC AAG TTC CTG GAG AAG CAG TAT GGC ACC TCA    1080

A  A  Y  A  N  F  L  I  G  A  V  N  L  P  A  A  A  L  G  M     380
GCA GCC TAT GCC AAC TTC CTC ATT GGT GCT GTG AAC CTC CCT GCT GCA GCC TTG GGG ATG    1140

L  F  G  G  I  L  M  K  R  F  V  F  S  L  Q  A  I  P  R  I     400
CTG TTT GGA GGA ATC CTC ATG AAG CGC TTT GTT TTC TCT CTA CAA GCC ATT CCC CGC ATA    1200

A  T  T  I  I  T  I  S  M  I  L  C  V  P  L  F  F  M  G  C     420
GCT ACC ACC ATC ATC ACC ATC TCC ATG ATC CTT TGT GTT CCT TTG TTC TTC ATG GGA TGC    1260
```

fchd502

FIG. 5A

7853-114

S T P T V A E V Y P P S T S S S I H P Q 440
TCC ACC CCA ACT GTG GCC GAA GTC TAC CCC CCT AGC ACA TCA AGT TCT ATA CAT CCG CAG 1320

S P A C R R D C S C P D S I F H P V C G 460
TCT CCT GCC TGC CGC AGG GAC TGC TCG TGC CCA GAT TCT ATC TTC CAC CCG GTC TGT GGA 1380

D N G I E Y L S P C H A G C S N I N M S 480
GAC AAT GGA ATC GAG TAC CTC TCC CCT TGC CAT GCC GGC TGC AGC AAC ATC AAC ATG AGC 1440

S A T S K Q L I Y L N C S C V T G G S A 500
TCT GCA ACC TCC AAG CAA CTG ATC TAT TTG AAC TGC AGC TGT GTG ACC GGG GGA TCC GCT 1500

S A K T G S C P V P C A H F L L P A I F 520
TCA GCA AAG ACA GGA TCG TGC CCT GTC CCC TGT GCC CAC TTC CTG CTC CCG GCC ATC TTC 1560

L I S F V S L I A C I S H N P L Y M M V 540
CTC ATC TCC TTC GTG TCC CTG ATA GCC TGC ATC TCC CAC AAC CCC CTC TAC ATG ATG GTT 1620

L R V V N Q E E K S F A I G V Q F L L M 560
CTG CGT GTG GTG AAC CAG GAG GAA AAG TCA TTT GCC ATC GGG GTG CAG TTC TTG TTG ATG 1680

R L L A W L P S P A L Y G L T I D H S C 580
CGC TTG CTG GCC TGG CTG CCA TCT CCA GCC CTC TAT GGC CTC ACC ATT GAC CAC TCC TGC 1740

I R W N S L C L G R R G A C A Y Y D N D 600
ATC CGG TGG AAC TCG CTG TGC TTG GGG AGG CGA GGG GCC TGC GCC TAC TAT GAC AAC GAT 1800

A L R D R Y L G L Q M G Y K A L G M L L 620
GCT CTC CGA GAC AGG TAC CTG GGC CTG CAG ATG GGC TAC AAG GCG CTG GGC ATG CTG CTG 1860

L C F I S W R V K K N K E Y N V Q K A A 640
CTT TGC TTC ATC AGC TGG AGG GTG AAG AAG AAC AAG GAG TAC AAC GTG CAG AAG GCG GCA 1920

G L I * 643
GGC CTC ATC TGA 1929

CCCCACCCTGGGCCACTGYCCTGCTCCAGAGAGTGGACCTTGACTCYTCCACACCTGCCTATACTCACTAATGTTAACA

CGTCATTTCCTKTTTGTATTTTTAAAMAAGA

FIG. 5B

fchd502

```
(chd523)
                        GGCTTACCATGGATGCGGCCGGGGATCCAGGGCTCAGAGGAGGACGCACCCGGCAGCCAG    61
CCGGGAACCTTCCCTGGCGGGCTCCCAGGGGGGGTCTCTTCCTCTCTCTAGCCCTGCTCAGGCATTCGGCAGGTCCAGC    140
AGAGGTACACCTCCTGCAGCGGGTTCCAAGTGCACCTCCAGCCTGATGGACCTGACCAAGGAGGCTTCCAGGAGCACAG    219
AAGGGGCTGCAACCCCAGGTACCCAGAGAGTGAGCAGCTCCACGCGGGACTGTGCACGGTGGCCGACACCCGCAGGGACG    298
CCCACCCGGACGAGCACGGGGAGGGCCCTGGCCTCCACGGATGCACCATGGGGGTGTGAGGAGCATCTGTTCTTCCCACT    377
CTCTGCAGTTAACAAACCCAACCCAAACCACCACAGGTGCTCCTCCTGGGGAGTTTCCTGTCTGACAAATGCCAGGCTC    456
ACTTCAAGGAGAATCACGCTTCTTTCTAAAGATGGATTCACCATTTAAAACAGAGCTCTGGAGCCTTTCGGCAAATCT    535
                         M   D   V   T   S   Q   A   R   G   V   G   L   E   M      14
TGAAAGCTGCACGGGGGCAGAGAC ATG GAT GTG ACT TCC CAA GCC CGG GGC GTA GGC CTG GAG ATG    600
 Y   P   G   T   A   Q   P   A   A   P   N   T   T   S   P   E   L   N   L   S      34
TAC CCA GGC ACC GCG CAG CCT GCG GCC CCC AAC ACC ACC TCC CCC GAG CTC AAC CTG TCC    664
 H   P   L   L   G   T   A   L   A   N   G   T   G   E   L   S   E   H   Q   Q      54
CAC CCG CTC CTG GGC ACC GCC CTG GCC AAT GGG ACA GGT GAG CTC TCG GAG CAC CAG CAA    724
 Y   V   I   G   L   F   L   S   C   L   Y   T   I   F   L   F   P   I   G   F      74
TAC GTG ATC GGC CTG TTC CTC TCG TGC CTC TAC ACC ATC TTC CTC TTC CCC ATC GGC TTT    780
 V   G   N   I   L   I   L   V   V   N   I   S   F   R   E   K   M   T   I   P      94
GTG GGC AAC ATC CTG ATC CTG GTG GTG AAC ATC AGC TTC CGC GAG AAG ATG ACC ATC CCC    840
 D   L   Y   F   I   N   L   A   V   A   D   L   I   L   V   A   D   S   L   I     114
GAC CTG TAC TTC ATC AAC CTG GCG GTG GCG GAC CTC ATC CTG GTG GCC GAC TCC CTC ATT    900
 E   V   F   N   L   H   E   R   Y   Y   D   I   A   V   L   C   T   F   M   S     134
GAG GTG TTC AAC CTG CAC GAG CGG TAC TAC GAC ATC GCC GTC CTG TGC ACC TTC ATG TCG    960
 L   F   L   R   V   N   M   Y   S   S   V   F   F   L   T   W   M   S   F   D     154
CTC TTC CTG CGG GTC AAC ATG TAC AGC AGC GTC TTC TTC CTC ACC TGG ATG AGC TTC GAC   1020
 R   Y   I   A   L   A   R   A   M   R   C   S   L   F   R   T   K   H   H   A     174
CGC TAC ATC GCC CTG GCC AGG GCC ATG CGC TGC AGC CTG TTC CGC ACC AAG CAC CAC GCC   1080
 R   L   S   C   G   L   I   W   M   A   S   V   S   A   T   L   V   P   F   T     194
CGG CTG AGC TGT GGC CTC ATC TGG ATG GCA TCC GTG TCA GCC ACG CTG GTG CCC TTC ACC   1140
 A   V   H   L   Q   H   T   D   E   A   C   F   C   F   A   D   V   R   E   V     214
GCC GTG CAC CTG CAG CAC ACC GAC GAG GCC TGC TTC TGT TTC GCG GAT GTC CGG GAG GTG   1200
 Q   W   L   E   V   T   L   G   F   I   V   P   F   A   I   I   G   L   C   Y     234
CAG TGG CTC GAG GTC ACG CTG GGC TTC ATC GTG CCC TTC GCC ATC ATC GGC CTG TGC TAC   1260
 S   L   I   V   R   V   L   V   R   A   H   R   H   R   G   L   R   P   R   R     254
TCC CTC ATT GTC CGG GTG CTG GTC AGG GCG CAC CGG CAC CGT GGG CTG CGG CCC CGG CGG   1320
 Q   K   A   L   R   M   I   L   A   V   V   L   V   F   F   V   C   W   L   P     274
CAG AAG GCG CTC CGC ATG ATC CTC GCA GTG GTG CTG GTC TTC TTC GTC TGC TGG CTG CCG   1380
 E   N   V   F   I   S   V   H   L   L   Q   R   T   Q   P   G   A   A   P   C     294
GAG AAC GTC TTC ATC AGC GTG CAC CTC CTG CAG CGG ACG CAG CCT GGG GCC GCT CCT TGC   1440
 K   Q   S   F   R   H   A   H   P   L   T   G   H   I   V   N   L   A   A   F     314
AAG CAG TCT TTC CGC CAT GCC CAC CCC CTC ACG GGC CAC ATT GTC AAC CTC GCC GCC TTC   1500
```

FIG. 6A

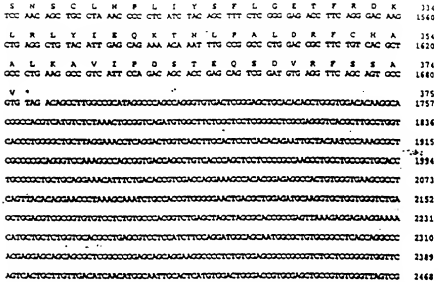

FIG. 6B

M A S P R A S R W P P P L L 14
CGGGCCCCGCGCCGTCACC ATG GCC TCG CCG CGC GCC TCG CGG TGG CCG CCG CCG CTC CTG 42

L L L L P L L L L P P A A P G T R D P P 34
CTG CTG TTG CTG CCG CTG CTG CTG CTG CCG CCG GCG GCC CCC GGG ACG CGG GAC CCG CCG 102

P S P A R R A L S L A P L A G A G L E L 54
CCT TCC CCG GCT CGC CGC GCG CTG AGC CTG GCG CCC CTC GCG GGA GCG GGG CTG GAG CTG 162

Q L E R R P E R E P P P T P P R E R R G 74
CAG CTG GAG CGC CGC CCG GAG CGC GAG CCG CCG CCC ACG CCG CCC CGG GAG CGC CGC GGG 222

P A T P G P S Y R A P E P G A A T Q R G 94
CCC GCG ACC CCC GGC CCC AGC TAC AGG GCC CCT GAG CCA GGC GCC GCG ACA CAG CGG GGA 282

P S G R A P R G G S A D A A W K H W P E 114
CCC TCC GGC CGG GCC CCC AGA GGC GGG AGC GCG GAT GCT GCC TGG AAA CAT TGG CCA GAA 342

S N T E A H V E N I T F Y Q N Q E D F S 134
AGT AAC ACT GAG GCC CAT GTA GAA AAC ATC ACC TTC TAT CAG AAT CAA GAG GAC TTT TCA 402

T V S S K E G V M V Q T S G K S H A A S 154
ACA GTG TCC TCC AAA GAG GGT GTG ATG GTT CAG ACC TCT GGG AAG AGC CAT GCT GCT TCG 462

D A P E N L T L L A E T A D A R G R S G 174
GAT GCT CCA GAA AAC CTC ACT CTA CTC GCT GAA ACA GCA GAT GCT AGA GGA AGG AGC GGC 522

S S S R T N F T I L P V G Y S L E I A T 194
TCT TCA AGT AGA ACA AAC TTC ACC ATT TTG CCT GTT GGG TAC TCA CTG GAG ATA GCA ACA 582

A L T S Q S G N L A S E S L H L P S S S 214
GCT CTG ACT TCC CAG AGT GGC AAC TTA GCC TCG GAA AGT CTT CAC CTG CCA TCC AGC AGT 642

S E F D E R I A A F Q T K S G T A S E M 234
TCA GAG TTC GAT GAA AGA ATT GCC GCT TTT CAA ACA AAG AGT GGA ACA GCC TCG GAG ATG 702

G T E R A M G L S E E W T V H S Q E A T 254
GGA ACA GAG AGG GCG ATG GGG CTG TCA GAA GAA TGG ACT GTG CAC AGC CAA GAG GCC ACC 762

T S A W S P S F L P A L E M G E L T T P 274
ACT TCG GCT TGG AGC CCG TCC TTT CTT CCT GCT TTG GAG ATG GGA GAG CTG ACC ACG CCT 822

S R K R N S S G P D L S W L H F Y R T A 294
TCT AGG AAG AGA AAT TCC TCA GGA CCA GAT CTC TCC TGG CTG CAT TTC TAC AGG ACA GCA 882

A S S P L L D L S S P S E S T E K L N N 314
GCT TCC TCT CCT CTC TTA GAC CTT TCC TCA CCT TCT GAA AGT ACA GAG AAG CTT AAC AAC 942

S T G L Q S S S V S Q T K T M H V A T V 334
TCC ACT GGC CTC CAG AGC TCC TCA GTC AGT CAA ACA AAG ACA ATG CAT GTT GCT ACC GTG 1002

F T D G G P R T L R S L T V S L G P V S 354
TTC ACT GAT GGT GGC CCG AGA ACG CTG CGA TCT TTG ACG GTC AGT CTG GGA CCT GTG AGC 1062

K T E G F P K D S R I A T T S S S V L L 374
AAG ACA GAA GGC TTC CCC AAG GAC TCC AGA ATT GCC ACG ACT TCA TCC TCA GTC CTT CTT 1122

S P S A V E S R R N S R V T G N P G D E 394
TCA CCC TCT GCA GTG GAA TCG AGA AGA AAC AGT AGA GTA ACT GGG AAT CCA GGG GAT GAG 1182

E F I E P S T E N E F G L T S L R W Q N 414
GAA TTC ATT GAA CCA TCC ACA GAA AAT GAA TTT GGA CTT ACG TCT TTG CGT TGG CAA AAT 1242

fchd528

FIG. 7A

```
D  S  P  T  F  G  E  H  Q  L  A  S  S  S  E  V  Q  N  G  S      434
GAT TCC CCA ACC TTT GGA GAA CAT CAG CTT GCC AGC AGC TCT GAG GTG CAA AAT GGA AGT  1302

P  M  S  Q  T  E  T  V  S  R  S  V  A  P  M  R  G  G  E  I      454
CCC ATG TCT CAG ACT GAG ACT GTG TCT AGG TCA GTC GCA CCC ATG AGA GGT GGA GAG ATC  1362

T  A  H  W  L  L  T  N  S  T  T  S  A  D  V  T  G  S  S  A      474
ACT GCA CAC TGG CTC TTG ACC AAC AGC ACA ACA TCT GCA GAT GTG ACA GGA AGC TCT GCT  1422

S  Y  P  E  G  V  N  A  S  V  L  T  Q  F  S  D  S  T  V  Q      494
TCA TAT CCT GAA GGT GTG AAT GCT TCA GTG TTG ACC CAG TTC TCA GAC TCT ACT GTA CAG  1482

S  G  G  S  H  T  A  L  G  D  R  S  Y  S  E  S  S  S  T  S      514
TCT GGA GGA AGT CAC ACA GCA TTG GGA GAT AGG AGT TAT TCA GAG TCT TCA TCT ACA TCT  1542

S  S  E  S  L  N  S  S  A  P  R  G  E  R  S  T  L  E  D  S      534
TCC TCG GAA AGC TTG AAT TCA TCA GCA CCA CGT GGA GAA CGT TCA ACC TTG GAA GAC AGC  1602

R  E  P  G  Q  A  L  G  D  S  S  A  N  A  E  D  R  T  S  G      554
CGA GAG CCA GGC CAA GCA CTA GGT GAC AGT TCC GCC AAT GCA GAG GAC AGG ACT TCT GGG  1662

V  P  S  L  G  T  H  T  L  A  T  V  T  G  N  G  E  R  T  L      574
GTG CCC TCT CTC GGC ACC CAC ACC TTG GCT ACT GTC ACT GGA AAC GGG GAA CGC ACA CTG  1722

R  S  V  T  L  T  N  T  S  M  S  T  T  S  G  E  A  G  S  P      594
CGG TCT GTC ACC CTC ACC AAC ACC AGC ATG AGC ACG ACT TCT GGG GAA GCA GGC AGC CCT  1782

A  A  A  M  P  Q  E  T  E  G  A  S  L  H  V  N  V  T  D  D      614
GCA GCG GCC ATG CCC CAA GAA ACA GAG GGT GCC TCT CTG CAC GTA AAC GTG ACG GAC GAC  1842

M  G  L  V  S  R  S  L  A  A  S  S  A  L  G  V  A  G  I  S      634
ATG GGC CTG GTC TCA CGG TCA CTG GCC GCC TCC AGT GCA CTC GGA GTC GCT GGG ATT AGC  1902

Y  G  Q  V  R  G  T  A  I  E  Q  R  T  S  S  D  H  T  D  H      654
TAC GGT CAA GTG CGT GGC ACA GCT ATT GAA CAA AGG ACT TCC AGC GAC CAC ACA GAC CAC  1962

T  Y  L  S  S  T  F  T  K  G  E  R  A  L  L  S  I  T  D  N      674
ACC TAC CTG TCA TCT ACT TTC ACC AAA GGA GAA CGG GCG TTA CTG TCC ATT ACA GAT AAC  2022

S  S  S  S  D  I  V  E  S  S  T  S  Y  I  K  I  S  N  S  S      694
AGT TCA TCC TCA GAC ATT GTG GAG AGC TCA ACT TCT TAT ATT AAA ATC TCA AAC TCT TCA  2082

H  S  E  Y  S  S  F  S  H  A  Q  T  E  R  S  N  I  S  S  Y      714
CAT TCA GAG TAT TCC TCC TTT TCT CAT GCT CAG ACT GAG AGA AGT AAC ATC TCA TCC TAT  2142

D  G  E  Y  A  Q  P  S  T  E  S  P  V  L  H  T  S  N  L  P      734
GAC GGG GAA TAT GCT CAG CCT TCT ACT GAG TCG CCA GTT CTG CAT ACA TCC AAC CTT CCG  2202

S  Y  T  P  T  I  N  M  P  N  T  S  V  V  L  D  T  D  A  E      754
TCC TAC ACA CCC ACC ATT AAT ATG CCG AAC ACT TCG GTT GTT CTG GAC ACT GAT GCT GAG  2262

F  V  S  D  S  S  S  S  S  S  S  S  S  S  S  S  S  G  P      774
TTT GTT AGT GAC TCC TCC TCC TCC TCT TCC TCC TCC TCC TCT TCT TCT TCT TCA GGG CCT  2322

P  L  P  L  P  S  V  S  Q  S  H  H  L  F  S  S  I  L  P  S      794
CCT TTG CCT CTG CCC TCT GTG TCA CAA TCC CAC CAT TTA TTT TCA TCA ATT TTA CCA TCA  2382

T  R  A  S  V  H  L  L  K  S  T  S  D  A  S  T  P  W  S  S      814
ACC AGG GCC TCT GTG CAT CTA CTA AAG TCT ACC TCT GAT GCA TCC ACA CCA TGG TCT TCC  2442

S  P  S  P  L  P  V  S  L  T  T  S  T  S  A  P  L  S  V  S      834
TCA CCA TCA CCT TTA CCA GTA TCC TTA ACG ACA TCT ACA TCT GCC CCA CTT TCT GTC TCA  2482
```

fchd528

FIG. 7B

```
Q  T  T  L  P  Q  S  S  S  T  P  V  L  P  R  A  R  E  T  P      854
CAA ACA ACC TTG CCA CAG TCA TCT TCT ACC CCT GTC CTG CCC AGG GCA AGG GAG ACT CCT  2542

V  T  S  F  Q  T  S  T  M  T  S  F  M  T  M  L  H  S  S  Q      874
GTG ACT TCA TTT CAG ACA TCA ACA ATG ACA TCA TTC ATG ACA ATG CTC CAT AGT AGT CAA  2602

T  A  D  L  K  S  Q  S  T  P  H  Q  E  K  V  I  T  E  S  K      894
ACT GCA GAC CTT AAG AGC CAG AGC ACC CCA CAC CAA GAG AAA GTC ATT ACA GAA TCA AAG  2682

S  P  S  L  V  S  L  P  T  E  S  T  K  A  V  T  T  N  S  P      914
TCA CCA AGC CTG GTG TCT CTG CCC ACA GAG TCC ACC AAA GCT GTA ACA ACA AAC TCT CCT  2742

L  P  P  S  L  T  E  S  S  T  E  Q  T  L  P  A  T  S  T  N      934
TTG CCT CCA TCC TTA ACA GAG TCC TCC ACA GAG CAA ACC CTT CCA GCC ACA AGC ACC AAC  2802

L  A  Q  M  S  P  T  F  T  T  T  I  L  K  T  S  Q  P  L  M      954
TTA GCA CAA ATG TCT CCA ACT TTC ACA ACT ACC ATT CTG AAG ACC TCT CAG CCT CTT ATG  2862

T  T  P  G  T  L  S  S  T  A  S  L  V  T  G  P  I  A  V  Q      974
ACC ACT CCT GGC ACC CTG TCA AGC ACA GCA TCT CTG GTC ACT GGC CCT ATA GCC GTA CAG  2922

T  T  A  G  K  Q  L  S  L  T  H  P  E  I  L  V  P  Q  I  S      994
ACT ACA GCT GGA AAA CAG CTC TCG CTG ACC CAT CCT GAA ATA CTA GTT CCT CAA ATC TCA  2982

T  E  G  G  I  S  T  E  R  N  R  V  I  V  D  A  T  T  G  L     1014
ACA GAA GGT GGC ATC AGC ACA GAA AGG AAC CGA GTG ATT GTG GAT GCT ACC ACT GGA TTG  3042

I  P  L  T  S  V  P  T  S  A  K  E  M  T  T  K  L  G  V  T     1034
ATC CCT TTG ACC AGT GTA CCC ACA TCA GCA AAA GAA ATG ACC ACA AAG CTT GGC GTT ACA  3102

A  E  Y  S  P  A  S  R  S  L  G  T  S  P  S  P  Q  T  T  V     1054
GCA GAG TAC AGC CCA GCT TCA CGT TCC CTC GGA ACA TCT CCT TCT CCC CAA ACC ACA GTT  3162

V  S  T  A  E  D  L  A  P  K  S  A  T  F  A  V  Q  S  S  T     1074
GTT TCC ACG GCT GAA GAC TTG GCT CCC AAA TCT GCC ACC TTT GCT GTT CAG AGC AGC ACA  3222

Q  S  P  T  T  L  S  S  S  A  S  V  N  S  C  A  V  N  P  C     1094
CAG TCA CCA ACA ACA CTG TCC TCT TCA GCC TCA GTC AAC AGC TGT GCT GTG AAC CCT TGT  3282

L  H  N  G  E  C  V  A  D  N  T  S  R  G  Y  H  C  R  C  P     1114
CTT CAC AAT GGC GAA TGC GTC GCA GAC AAC ACC AGC CGT GGC TAC CAC TGC AGG TGC CCG  3342

P  S  W  Q  G  D  D  C  S  V  D  V  N  E  C  L  S  N  P  C     1134
CCT TCC TGG CAA GGG GAT GAT TGC AGT GTG GAT GTG AAT GAG TGC CTG TCG AAC CCC TGC  3402

P  S  T  A  T  C  N  N  T  Q  G  S  F  I  C  K  C  P  V  G     1154
CCA TCC ACA GCC ACG TGC AAC AAT ACT CAG GGA TCC TTT ATC TGC AAA TGC CCG GTT GGG  3462

Y  Q  L  E  K  G  I  C  N  L  V  R  T  F  V  T  E  F  K  L     1174
TAC CAG TTG GAA AAA GGG ATA TGC AAT TTG GTT AGA ACC TTC GTG ACA GAG TTT AAA TTA  3522

K  R  T  F  L  N  T  T  V  E  K  H  S  D  L  Q  E  V  E  N     1194
AAG AGA ACT TTT CTT AAT ACA ACT GTG GAA AAA CAT TCA GAC CTA CAA GAA GTT GAA AAT  3582

E  I  T  K  T  L  N  M  C  F  S  A  L  P  S  Y  I  R  S  T     1214
GAG ATC ACC AAA ACG TTA AAT ATG TGT TTT TCA GCG TTA CCT AGT TAC ATC CGA TCT ACA  3642

V  H  A  S  R  E  S  N  A  V  V  I  S  L  Q  T  T  F  S  L     1234
GTT CAC GCC TCT AGG GAG TCC AAC GCG GTG GTG ATC TCA CTG CAA ACA ACC TTT TCC CTG  3702

A  S  N  V  T  L  F  D  L  A  D  R  M  Q  K  C  V  N  S  C     1254
GCC TCC AAT GTG ACG CTA TTT GAC CTG GCT GAT AGG ATG CAG AAA TGT GTC AAC TCC TGC  3762
```

fchd528

FIG. 7C

```
K  S  S  A  E  V  C  Q  L  L  G  S  Q  R  R  I  F  R  A  G    1274
AAG TCC TCT GCT GAG GTC TGC CAG CTC TTG GGA TCT CAG AGG CGG ATC TTT AGA GCG GGC    3822

S  L  C  K  R  K  S  P  E  C  D  K  D  T  S  I  C  T  D  L    1294
AGC TTG TGC AAG CGG AAG AGT CCC GAA TGT GAC AAA GAC ACC TCC ATC TGC ACT GAC CTG    3882

D  G  V  A  L  C  Q  C  K  S  G  Y  F  Q  F  N  K  M  D  H    1314
GAC GGC GTT GCC CTG TGC CAG TGC AAG TCG GGA TAC TTT CAG TTC AAC AAG ATG GAC CAC    3942

S  C  R  A  C  E  D  G  Y  R  L  E  N  E  T  C  M  S  C  P    1334
TCC TGC CGA GCA TGT GAA GAT GGA TAT AGG CTT GAA AAT GAA ACC TGC ATG AGT TGC CCA    4002

F  G  L  G  G  L  N  C  G  N  P  Y  Q  L  I  T  V  V  I  A    1354
TTT GGC CTT GGT GGT CTC AAC TGT GGA AAC CCC TAT CAG CTT ATC ACT GTG GTG ATC GCA    4062

A  A  G  G  G  L  L  L  I  L  G  I  A  L  I  V  T  C  C  R    1374
GCC GCG GGA GGT GGG CTC CTG CTC ATC CTA GGC ATC GCA CTG ATT GTT ACC TGT TGC AGA    4122

K  N  K  N  D  I  S  K  L  I  F  K  S  G  D  F  Q  M  S  P    1394
AAG AAT AAA AAT GAC ATA AGC AAA CTC ATC TTC AAA AGT GGA GAT TTC CAA ATG TCC CCA    4182

Y  A  E  Y  P  K  N  P  R  S  Q  E  W  G  R  E  A  I  E  M    1414
TAT GCT GAA TAC CCC AAA AAT CCT CGC TCA CAA GAA TGG GGC CGA GAA GCT ATT GAA ATG    4242

H  E  N  G  S  T  K  N  L  L  Q  M  T  D  V  Y  Y  S  P  T    1434
CAT GAG AAT GGA AGT ACC AAA AAC CTC CTC CAG ATG ACG GAT GTG TAC TAC TCG CCT ACA    4302

S  V  R  N  P  E  L  E  R  N  G  L  Y  P  A  Y  T  G  L  P    1454
AGT GTA AGG AAT CCA GAA CTT GAA CGA AAC GGA CTC TAC CCG GCC TAC ACT GGA CTG CCA    4362

G  S  R  H  S  C  I  F  P  G  Q  Y  N  P  S  F  I  S  D  E    1474
GGA TCA CGG CAT TCT TGC ATT TTC CCC GGA CAG TAT AAC CCG TCT TTC ATC AGT GAT GAA    4422

S  R  R  R  D  Y  F  *                                        1481
AGC AGA AGA AGA GAC TAC TTT TAA  GTCCAGGAGAGAGAGGGACTCATTGCTCTGAGCCAG           4482

  TCACCTGGGACCTCTGCTCAGAGGACCGCACCAGGAGGCTGCGCCCAGGATTTGTCGGGA                  4542

  GCCACGCTGAGTGGCAAGCAGGAAGAGGGACAGGCATGCGGGGCGTGACCACAGTGGAGG                  4602

  AGACAGGTGGATGTGGAACCACAGGCTGCTCATTCAGCACCTTTGTTGTTACTGTGAACG                  4662

  TGAATGTGGGCCAGTATCAAGAGAGTCTCTCTGAGTGACTGCACCATGGCACTGGCACCA                  4722

  GGGCGACTATTAGCCAGGGCAGACCACTAGACTTCAGTGCAGGGACCTGGTTTTCCCTTC                  4782

  GTTTGCACTTTAGTAAATTGGGTGGGAGGTTTCCTTTTGGATCTGTTTTGAGACTGTTCC                  4842

  AGAAAGAAGGCTTCCTTTCCCGAGACACTTCCATAGGCAGCAATTTGGTGATTCATTTGC                  4902

  ASCAAAATACTGGCTGTTAATTATTTTCCTGCCCAGCRCCTGCGTGCTAAACAACAGAT                   4962

  GAGGATGASCGTACCACTGAAGTCTGAAGATGTCGCCATTGAACGGACAGTGTTTTTCATA                 5022

  TGTTTCTAGGTTGTCTTATGCTACAGTTTCCAAGCCASCCCCCACAGTGAGGAAATGTGT                  5082

  GAGGCACCGCACACAACTGCAATGTGTTTYTTAAGTCAAGGTGACACATGTATTTAAGAT                  5142

  TTTTTTTTAAAATCTCYTTGCAGTTAAATCTCACTTTYTCAAACAAGCCTGGATCAGGGC                  5202

  AAAACAACTTATATTYGGTTTTAGCTGGAGGCTCAGCAGGCAGATTGCAGGCAGGGGGC                  5262

  ACTTTTCATCCATGAGGGCCCAGCCTGGGGCCTGGGACTCTGATCACCATTGTGGAGGCC                  5322
```

fchd528

FIG. 7D

7853-114

AGAGGCAMCTGCGTATGGAGGAGAAATGTCAAACTGAACGCAGGTTTCACCACTCTAGGA 5382

AAGCAGCTTGTTGACCCCCTGCASCTGGATGTGGTTAGAGGGATGGGCTGAATAGSCAGG 5442

TTAGATTTCCTGCATCAACAGTGCTTTGGGAASCTGTGTGGATTCCTGAGGAAGAACAGG 5502

GAGCCGAGATGGAGCCACACATGAATTYGCTCACCGGCTACTGCAGCACTTTGTACCCAG 5562

AATCTCATGTCCACAAACCCCATGTAAACTTTCAACCACTCAAAGSTGTTTATTCGGCTG 5622

AAGAAATAACTTTTKTTTCTCACCCAGTCATTTGTACCTCTTCATATGGSTATGTCGCAC 5682

CCTCCAGAAACGTGGTTATACTKCCAGTCAGTGTGGGAGAACTGAAGACTTCCGGTTGGT 5742

CGAGGAACTGAGGGTTGACCTTCGGGAAGGAAGTTCCACTCATCTTATTTATTATGCCTG 5802

TGATGTGGGTCCTGCCAGGGAGACATCCAGTACTCGGTGTCTKTAATTGCCACCTGGGGA 5862

ACTGTGTTTATTGGCCTTCTTTGGGGCATCCTGGKTTCGGATGAAGTGAGGGGAATACAG 5922

AGGTAAAAGAATTGTCTCCACCCTGAAGCGGGGAGTCCCGCTTCACATTTCTGGAAATGG 5982

TGCAGCCACTGGGGACAGTTCTGCCCCGGGCATGGTTGTTTCTTCAAGGTCCTCTAAATA 6042

TAATCCCTATTCTTACATAATCCTTGGCCCTGATGGTTTTAAGCAAGAACTCCTGTGTCC 6102

MATGGTCTCCACCACTCACCATCACCCTGCTGTAGCAAGAGTCCTAGTCAGGGGAGGTGC 6162

ATTTTAGTAGTTACATTGCACTTATCCATGAGATAAATAAAAGGAGAVCTGTTTTTATCA 6222

GTGGAGGCTAACCTAAAATTTCAAAGTGTCGCCTTTTTGAAATCTTGGGCCTCTCTCTCT 6282

GTAGAACCAATGCCCCTTTGTGGCTCACGGCCTCGCACCTAACTGGAGAGTTCTGAGCTC 6342

CTGCAGCTCACCTGAGCCCACAGACTAGGCTTCTTGGCTCCTTCCGC 6389

fchd528

FIG. 7E

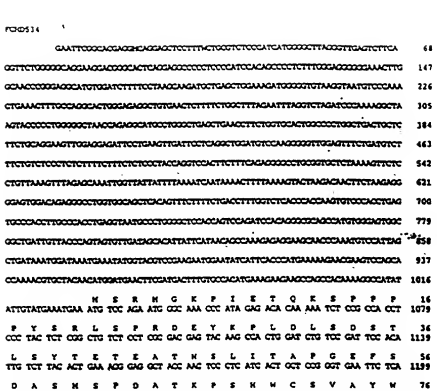

FIG. 8A

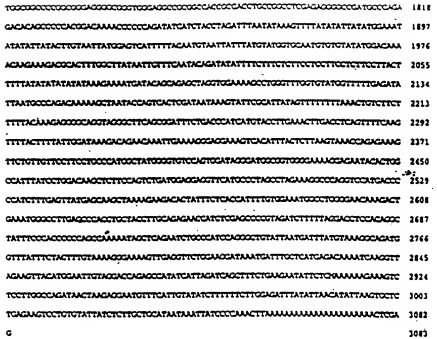

FIG. 8B

```
ACGAGGACGACAGGCTGTGCGCGGTCTGCACGGCGCTCCGCGGCGGAGCTTCATGTGGGGCTGCGACCCGCGCAGCCGG  79

                                                                          M     1
CGCCTCGCTGAGGGAACGGACCCCCGGTAACCGGAGACCGCCTTCCCCCCCACCCCTGGCGCCAAAGGATATCGT ATG   157

 F  R  S  K  R  S  G  L  V  R  R  L  W  R  S  R  V  V  P  D                    21
TTC AGG TCC AAA CGC TCG GGG CTG GTG CGG CGA CTT TGG CGA AGT CGT GTG GTC CCC GAC   217

 R  E  E  G  G  S  G  G  G  G  G  G  D  E  D  G  S  L  G  S                    41
CGG GAG GAA GGC GGC AGC GGC GGC GGC GGT GGC GGC GAC GAG GAT GGG AGC TTG GGC AGC   277

 R  A  E  P  A  P  R  A  R  E  G  G  G  C  G  R  S  E  V  R                    61
CGA GCT GAG CCG GCC CCG CGG GCA AGA GAG GGC GGA GGC TGC GGC CGC TCC GAA GTC CGC   337

 P  V  A  P  R  R  P  R  D  A  V  G  Q  R  G  A  Q  G  A  G                    81
CCG GTA GCC CCG CGG CGG CCC CGG GAC GCA GTG GGA CAG CGA GGC GCC CAG GGC GCG GGG   397

 R  R  R  R  A  G  G  P  P  R  P  M  S  E  P  G  A  G  A  G                   101
AGG CGC CGG CGC GCA GGG GGC CCC CCG AGG CCC ATG TCG GAG CCA GGG GCC GGC GCT GGG   457

 S  S  L  L  D  V  A  E  P  G  G  P  G  W  L  P  E  S  D  C                   121
AGC TCC CTG CTG GAC GTG GCG GAG CCG GGA GGC CCG GGC TGG CTG CCC GAG AGT GAC TGC   517

 E  T  V  T  C  C  L  F  S  E  R  D  A  A  G  A  P  R  D  A                   141
GAG ACG GTG ACC TGC TGT CTC TTT TCG GAG CGG GAC GCC GCC GGC GCG CCC CGG GAC GCC   577

 S  D  P  L  A  G  A  A  L  E  P  A  G  G  G  R  S  R  E  A                   161
AGC GAC CCC CTG GCC GGG GCG GCC CTG GAG CCG GCG GGC GGC GGG CGG AGT CGC GAA GCG   637

 R  S  R  L  L  L  L  E  Q  E  L  K  T  V  T  Y  S  L  L  K                   181
CGC TCG CGG CTC CTG CTG CTG CTG GAG CAG GAA CTC AAA ACC GTC ACG TAC TCG CTG CTG AAG   697

 R  L  K  E  R  S  L  D  T  L  L  E  A  V  E  S  R  G  G  V                   201
CGG CTC AAG GAG CGC TCG CTG GAC ACG CTG CTG GAG GCG GTG GAG TCC CGC GGC GGC GTG   757

 P  G  G  C  V  L  V  P  R  A  D  L  R  L  G  G  Q  P  A  P                   221
CCG GGC GGC TGC GTG CTG GTG CCG CGC GCC GAC CTC CGC CTG GGC GGC CAG CCC GCG CCG   817

 P  Q  L  L  L  G  R  L  F  R  W  P  D  L  Q  H  A  V  E  L                   241
CCG CAG CTG CTG CTC GGC CGC CTC TTT CGC TGG CCC GAC CTG CAG CAC GCC GTG GAG CTG   877

 K  P  L  C  G  C  H  S  F  A  A  A  A  D  G  P  T  V  C  C                   261
AAG CCC CTG TGC GGC TGC CAC AGC TTC GCC GCC GCC GCC GAC GGC CCT ACC GTG TGC TGC   937

                                                  => Same as short
 N  P  Y  H  F  S  R  L  C  G  P  E  S  P  P  P  P  Y  S  R                   281
AAC CCC TAC CAC TTC AGC CGG CTC TGC GGG CCC GAA TCT CCG CCA CCT CCC TAC TCT CGG   997

 L  S  P  R  D  E  Y  K  P  L  D  L  S  D  S  T  L  S  Y  T                   301
CTG TCT CCT CGC GAC GAG TAC AAG CCA CTG GAT CTG TCC GAT TCC ACA TTG TCT TAC ACT  1057

 E  T  E  A  T  N  S  L  I  T  A  P  G  E  F  S  D  A  S  M                   321
GAA ACG GAG GCT ACC AAC TCC CTC ATC ACT GCT CCG GGT GAA TTC TCA GAC GCC AGC ATG  1117

                    Met ->
 S  P  D  A  T  K  P  S  H  W  C  S  V  A  Y  W  E  H  R  T                   341
TCT CCG GAC GCC ACC AAG CCG AGC CAC TGG TGC AGC GTG GCG TAC TGG GAG CAC CGG ACG  1177
                      Met

 R  V  G  R  L  Y  A  V  Y  D  Q  A  V  S  I  F  Y  D  L  P                   361
CGC GTG GGC CGC CTC TAT GCG GTG TAC GAC CAG GCC GTC AGC ATC TTC TAC GAC CTA CCT  1237
```

Tong

FIG. 9A

```
 Q   G   S   G   F   C   L   G   Q   L   N   L   E   Q   R   S   E   S   V   R    381
CAG GGC AGC GGC TTC TGC CTG GGC CAG CTC AAC CTG GAG CAG CGC AGC GAG TCG GTG CGG  1297

 R   T   R   S   K   I   G   F   G   I   L   L   S   K   E   P   D   G   V   W    401
CGA ACG CGC AGC AAG ATC GGC TTC GGC ATC CTG CTC AGC AAG GAG CCC GAC GGC GTG TGG  1357

 A   Y   N   R   G   E   H   P   I   F   V   N   S   P   T   L   D   A   P   G    421
GCC TAC AAC CGC GGC GAG CAC CCC ATC TTC GTC AAC TCC CCG ACG CTG GAC GCG CCC GGC  1417

 G   R   A   L   V   V   R   K   V   P   P   G   Y   S   I   K   V   F   D   F    441
GGC CGC GCC CTG GTC GTG CGC AAG GTG CCC CCC GGC TAC TCC ATC AAG GTG TTC GAC TTC  1477

 E   R   S   G   L   Q   H   A   P   E   P   D   A   A   D   G   P   Y   D   P    461
GAG CGC TCG GGC CTG CAG CAC GCG CCC GAG CCC GAC GCC GCC GAC GGC CCC TAC GAC CCC  1537

 N   S   V   R   I   S   F   A   K   G   W   G   P   C   Y   S   R   Q   F   I    481
AAC AGC GTC CGC ATC AGC TTC GCC AAG GGC TGG GGG CCC TGC TAC TCC CGG CAG TTC ATC  1597

 T   S   C   P   C   W   L   E   I   L   L   N   N   P   R   *                    497
ACC TCC TGC CCC TGC TGG CTG GAG ATC CTC CTC AAC AAC CCC AGA TAG                  1645

TGGCGGCCCCGGCGGGAGGGGCGGGTGGGAGGCCGCGGCCACCGCCACCTGCCGGCCTCGAGAGGGGCCGATGCCCAGA  1724

GACACAGCCCCCACGGACAAAACCCCCCAGATATCATCTACCTAGATTTAATATAAAGTTTTATATATTATATGGAAAA  1803

AAAAAAAAAAAAAA                                                                   1817
```

FIG. 9B

GCACGAGCGGAGAGCCGCGCAGGGCGCGGGCCGCGCGGGGTGGGGCAGCCGGAGCGCA

GGCCCCCGATCCCCGGCGGGCGCCCCCGGGCCCCCGCGCGCGCCCCGGCCTCCGGGAGACTGGCGCATGCCACGGAGCG

CCCCTCGGGCCGCCGCCGCTCCTGCCCGGGCCCCTGCTGCTGCTGCTGTCGCCTGCGCCTGCTGCCCCAACTCGGCGCC

M F 2
CGACTCACAAAGAAACATCATGTTGGCTCCTTAGCAGGCAAACGACTTTTCTCCTCGCCTCCTCGCCCCGC ATG TTC 6

R T K R S A L V R R L W R S R A P G G E 22
AGG ACC AAA CGA TCT GCG CTC GTC CGG CGT CTC TGG AGG AGC CGT GCG CCC GGC GGC GAG 66

D E E E G A G G G G G G G E L R G E G A 42
GAC GAG GAG GAG GGC GCA GGG GGA GGT GGA GGA GGA GGC GAG CTG CGG GGA GAA GGG GCG 126

T D S R A H G A G G G G P G R A G C C L 62
ACG GAC AGC CGA GCG CAT GGG GCC GGT GGC GGC GGC CCG GGC AGG GCT GGA TGC TGC CTG 186

G K A V R G A K G H H H P H P P A A G A 82
GGC AAG GCG GTG CGA GGT GCC AAA GGT CAC CAC CAT CCC CAC CCG CCA GCC GCG GGC GCC 246

G A A G G A E A D L K A L T H S V L K K 102
GGC GCG GCC GGG GGC GCC GAG GCG GAT CTG AAG GCG CTC ACG CAC TCG GTG CTC AAG AAA 306

L K E R Q L E L L L Q A V E S R G G T R 122
CTG AAG GAG CGG CAG CTG GAG CTG CTG CTC CAG GCC GTG GAG TCC CGC GGC GGG ACG CGC 366

T A C L L L P G R L D C R L G P G A P A 142
ACC GCG TGC CTC CTG CTG CCC GGC CGC CTG GAC TGC AGG CTG GGC CCG GGG GCG CCC GCC 426

G A Q P A Q P P S S Y S L P L L L C K V 162
GGC GCG CAG CCT GCG CAG CCG CCC TCG TCC TAC TCG CTC CCC CTC CTG CTG TGC AAA GTG 486

F R W P D L R H S S E V K R L G C C E S 182
TTC AGG TGG CCG GAT CTC AGG CAT TCC TCG GAA GTC AAG AGG CTG TGT TGC TGT GAA TCT 546

Y G K I N P E L V C C N P H H L S R L C 202
TAC GGG AAG ATC AAC CCC GAG CTG GTG TGC TGC AAC CCC CAT CAC CTT AGC CGA CTC TGC 606

E L E S P P P P Y S R Y P M D F L K P T 222
GAA CTA GAG TCT CCC CCC CCT CCT TAC TCC AGA TAC CCG ATG GAT TTT CTC AAA CCA ACT 666

A D C P D A V P S S A E T G G T N Y L A 242
GCA GAC TGT CCA GAT GCT GTG CCT TCC TCC GCT GAA ACA GGG GGA ACG AAT TAT CTG GCC 726

P G G L S D S Q L L L E P G D R S H W C 262
CCT GGG GGG CTT TCA GAT TCC CAA CTT CTT CTG GAG CCT GGG GAT CGG TCA CAC TGG TGC 786

V V A Y W E E K T R V G R L Y C V Q E P 282
GTG GTG GCA TAC TGG GAG GAG AAG ACG AGA GTG GGG AGG CTC TAC TGT GTC CAG GAG CCC 846

S L D I F Y D L P Q G N G F C L G Q L N 302
TCT CTG GAT ATC TTC TAT GAT CTA CCT CAG GGG AAT GGC TTT TGC CTC GGA CAG CTC AAT 906

S D N K S Q L V Q K V R S K I G C G I Q 322
TCG GAC AAC AAG AGT CAG CTG GTG CAG AAG GTG CGG AGC AAA ATC GGC TGC GGC ATC CAG 966

L T R E V D G V W V Y N R S S Y P I F I 342
CTG ACG CGG GAG GTG GAT GGT GTG TGG GTG TAC AAC CGC AGC AGT TAC CCC ATC TTC ATC 1026

K S A T L D N P D S R T L L V H K V F P 362
AAG TCC GCC ACA CTG GAC AAC CCG GAC TCC AGG ACG CTG TTG GTA CAC AAG GTG TTC CCC 1086

FIG. 10A

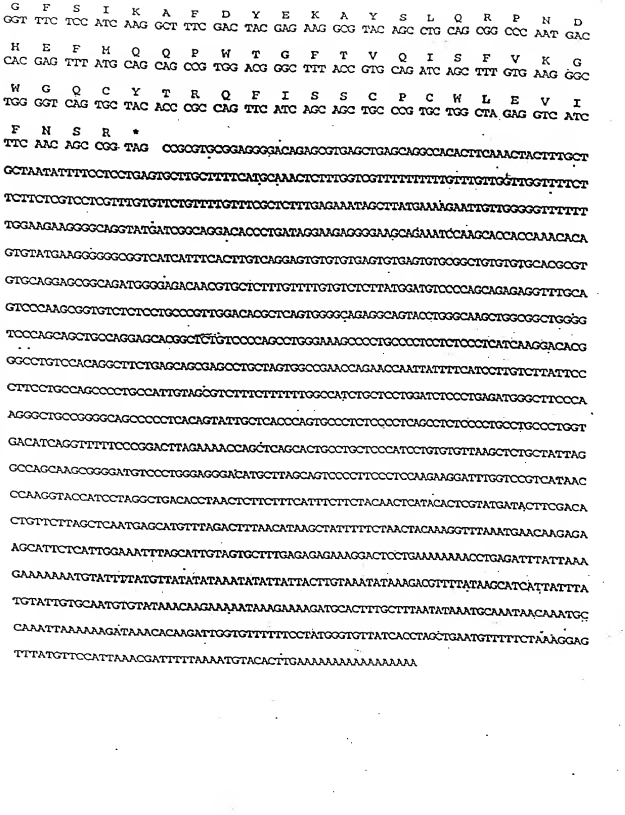
G F S I K A F D Y E K A Y S L Q R P N D
GGT TTC TCC ATC AAG GCT TTC GAC TAC GAG AAG GCG TAC AGC CTG CAG CGG CCC AAT GAC
382
1146
H E F M Q Q P W T G F T V Q I S F V K G
CAC GAG TTT ATG CAG CAG CCG TGG ACG GGC TTT ACC GTG CAG ATC AGC TTT GTG AAG GGC
402
1206
W G Q C Y T R Q F I S S C P C W L E V I
TGG GGT CAG TGC TAC ACC CGC CAG TTC ATC AGC AGC TGC CCG TGC TGG CTA GAG GTC ATC
422
1266
F N S R *
TTC AAC AGC CGG TAG CCGCGTGCGGAGGGGACAGAGCGTGAGCTGAGCAGGCCACACTTCAAACTACTTTGCT
426
1278
GCTAATATTTTCCTCCTGAGTGCTTGCTTTTCATGCAAACTCTTTGGTCGTTTTTTTTTTGTTTGTTGTTGGTTTTCT
TCTTCTCGTCCTCGTTTGTTGTTCTGTTTTGTTTCGCTCTTTGAGAAATAGCTTATGAAAGAATTGTTGGGGGTTTTTT
TGGAAGAAGGGGCAGGTATGATCGGCAGGACACCCTGATAGGAAGAGGGAAGCAGAAATCCAAGCACCACCAAACACA
GTGTATGAAGGGGGGCGGTCATCATTTCACTTGTCAGGAGTGTGTGTGAGTGTGAGTGTGCGGCTGTGTGTGCACGCGT
GTGCAGGAGCGGCAGATGGGGAGACAACGTGCTCTTTGTTTTGTGTCTCTTATGGATGTCCCCAGCAGAGAGGTTTGCA
GTCCCAAGCGGTGTCTCTCCTGCCCGTTGGACACGCTCAGTGGGGCAGAGGCAGTACCTGGGCAAGCTGGCGGCTGGGG
TCCCAGCAGCTGCCAGGAGCAGGCTCTGTCCCCAGCCTGGGAAAGCCCCTGCCCCTCCTCTCCCTCATCAAGGACACG
GGCCTGTCCACAGGCTTCTGAGCAGCGAGCCTGCTAGTGGCCGAACCAGAACCAATTATTTCATCCTTGTCTTATTCC
CTTCCTGCCAGCCCCTGCCATTGTAGCGTCTTTCTTTTTTGGCCATCTGCTCCTGGATCTCCCTGAGATGGGCTTCCCA
AGGGCTGCCGGGGCAGCCCCCTCACAGTATTGCTCACCCAGTGCCCTCTCCCCTCAGCCTCTCCCCTGCCTGCCCTGGT
GACATCAGGTTTTTCCCGGACTTAGAAAACCAGCTCAGCACTGCCTGCTCCCATCCTGTGTGTTAAGCTCTGCTATTAG
GCCAGCAAGCGGGGATGTCCCTGGGAGGGACATGCTTAGCAGTCCCCTTCCCTCCAAGAAGGATTTGGTCCGTCATAAC
CCAAGGTACCATCCTAGGCTGACACCTAACTCTTCTTTCATTTCTTCTACAACTCATACACTCGTATGATACTTCGACA
CTGTTCTTAGCTCAATGAGCATGTTAGACTTTAACATAAGCTATTTTTCTAACTACAAAGGTTTAAATGAACAAGAGA
AGCATTCTCATTGGAAATTTAGCATTGTAGTGCTTTGAGAGAGAAAGGACTCCTGAAAAAAAACCTGAGATTTATTAAA
GAAAAAAATGTATTTTATGTTATATATAAATATATTATTACTTGTAAATATAAAGACGTTTTATAAGCATCATTATTTA
TGTATTGTGCAATGTGTATAAACAAGAAAAATAAAGAAAAGATGCACTTTGCTTTAATATAAATGCAAATAACAAATGC
CAAATTAAAAAAGATAAACACAAGATTGGTGTTTTTTCCTATGGGTGTTATCACCTAGCTGAATGTTTTTCTAAAGGAG
TTTATGTTCCATTAAACGATTTTTAAAATGTACACTTGAAAAAAAAAAAAAAAAA

FIG. 10B

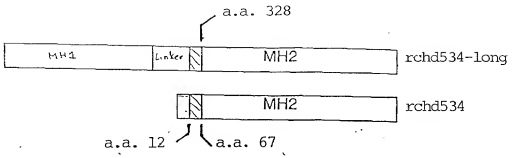

FIG. 11

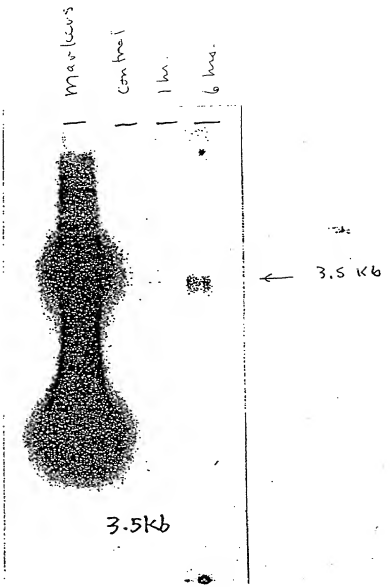

FIG. 12

DECLARATION AND POWER OF ATTORNEY

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below at 201 et seq. underneath my name.

I believe I am the original, first and sole inventor if only one name is listed at 201 below, or an original, first and joint inventor if plural names are listed at 201 et seq. below, of the subject matter which is claimed and for which a patent is sought on the invention entitled

COMPOSITIONS AND METHODS FOR THE TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE

and for which a patent application:
☒ is attached hereto
☐ was filed in the United States on _____ as Application No. _____ *(for declaration not accompanying application)*
with amendment(s) filed on _____ *(if applicable)*
☐ was filed as PCT international Application No. _____ on _____ and was amended under PCT Article 19 on _____ *(if applicable)*

I hereby state that I have reviewed and understand the contents of the above identified application, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, §1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, §119(a)-(d) of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed:

EARLIEST FOREIGN APPLICATION(S), IF ANY, FILED PRIOR TO THE FILING DATE OF THE APPLICATION			
APPLICATION NUMBER	COUNTRY	DATE OF FILING (day, month, year)	PRIORITY CLAIMED
			YES ☐ NO ☐

I hereby claim the benefit under Title 35, United States Code, §119(e) of any United States provisional application(s) listed below.

APPLICATION NUMBER	FILING DATE
60/011,787	February 16, 1996

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code §112, I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations, §1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application:

APPLICATION SERIAL NO.	FILING DATE	STATUS		
		PATENTED	PENDING	ABANDONED
08//870,434	June 6, 1997		x	
08/799,910	February 13, 1997		x	
08/599,654	February 9, 1996		x	
08/485,573	June 7, 1995		x	
08/386,844	February 10, 1995		x	

POWER OF ATTORNEY: As a named inventor, I hereby appoint S. Leslie Misrock (Reg. No. 18872), Harry C. Jones, III (Reg. No. 20280), Berj A. Terzian (Reg. No. 20060), Gerald J. Flintoft (Reg. No. 20823), David Weild, III (Reg. No. 21094), Jonathan A. Marshall (Reg. No. 24614), Barry D. Rein (Reg. No. 22411), Stanton T. Lawrence, III (Reg. No. 25736), Isaac Jarkovsky (Reg. No. 22713), Joseph V. Colaianni (Reg. No. 20019), Charles E. McKenney (Reg. No. 22795), Philip T. Shannon (Reg. No. 24278), Francis E. Morris (Reg. No. 24615), Charles E. Miller (Reg. No. 24576), Gidon D. Stern (Reg. No. 27469), John J. Lauter, Jr. (Reg. No. 27814), Brian M. Poissant (Reg. No. 28462), Brian D. Coggio (Reg. No. 27624), Rory J. Radding (Reg. No. 28749), Stephen J. Harbulak (Reg. No. 29166), Donald J. Goodell (Reg. No. 19766), James N. Palik (Reg. No. 25510), Thomas E. Friebel (Reg. No. 29258), Laura A. Coruzzi (Reg. No. 30742), Jennifer Gordon (Reg. No. 30753), Jon R. Stark (Reg. No. 30111), Allan A. Fanucci (Reg. No. 30256), Geraldine F. Baldwin (Reg. No. 31232), Victor N. Balancia (Reg. No. 31231), Samuel B. Abrams (Reg. No. 30605), Steven I. Wallach (Reg. No. 35402), Marcia H. Sundeen (Reg. No. 30893), Paul J. Zegger (Reg. No. 33821), Edmond R. Bannon (Reg. No. 32110), Bruce J. Barker (Reg. No. 33291), Adriane M. Antler (Reg. No. 32605), Ann L. Gisolfi (Reg. No. 31956), Mark A. Farley (Reg. No. 33170), and James G. Markey (Reg. No. 31636), all of Pennie & Edmonds LLP, whose addresses are 1155 Avenue of the Americas, New York, New York 10036, 1667 K Street N.W., Washington, DC 20006 and 3300 Hillview Avenue, Palo Alto, CA 94304, and each of them, my attorneys, to prosecute this application, and to transact all business in the Patent and Trademark Office connected therewith.

SEND CORRESPONDENCE TO:	PENNIE & EDMONDS LLP 1155 AVENUE OF THE AMERICAS NEW YORK, NEW YORK 10036	DIRECT TELEPHONE CALLS TO: PENNIE & EDMONDS LLP DOCKETING (212) 790-9090

201	FULL NAME OF INVENTOR	LAST NAME FALB	FIRST NAME Dean	MIDDLE NAME A.	
	RESIDENCE & CITIZENSHIP	CITY Wellesley	STATE OR FOREIGN COUNTRY MA	COUNTRY OF CITIZENSHIP U.S.A.	
	POST OFFICE ADDRESS	STREET 28 Swarthmore Road	CITY Wellesley	STATE OR COUNTRY MA	ZIP CODE 02181
202	FULL NAME OF INVENTOR	LAST NAME	FIRST NAME	MIDDLE NAME	
	RESIDENCE & CITIZENSHIP	CITY	STATE OR FOREIGN COUNTRY	COUNTRY OF CITIZENSHIP	
	POST OFFICE ADDRESS	STREET	CITY	STATE OR COUNTRY	ZIP CODE
203	FULL NAME OF INVENTOR	LAST NAME	FIRST NAME	MIDDLE NAME	
	RESIDENCE & CITIZENSHIP	CITY	STATE OR FOREIGN COUNTRY	COUNTRY OF CITIZENSHIP	
	POST OFFICE ADDRESS	STREET	CITY	STATE OR COUNTRY	ZIP CODE
204	FULL NAME OF INVENTOR	LAST NAME	FIRST NAME	MIDDLE NAME	
	RESIDENCE & CITIZENSHIP	CITY	STATE OR FOREIGN COUNTRY	COUNTRY OF CITIZENSHIP	
	POST OFFICE ADDRESS	STREET	CITY	STATE OR COUNTRY	ZIP CODE
205	FULL NAME OF INVENTOR	LAST NAME	FIRST NAME	MIDDLE NAME	
	RESIDENCE & CITIZENSHIP	CITY	STATE OR FOREIGN COUNTRY	COUNTRY OF CITIZENSHIP	
	POST OFFICE ADDRESS	STREET	CITY	STATE OR COUNTRY	ZIP CODE
206	FULL NAME OF INVENTOR	LAST NAME	FIRST NAME	MIDDLE NAME	
	RESIDENCE & CITIZENSHIP	CITY	STATE OR FOREIGN COUNTRY	COUNTRY OF CITIZENSHIP	
	POST OFFICE ADDRESS	STREET	CITY	STATE OR COUNTRY	ZIP CODE

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

SIGNATURE OF DEAN A. FALB	SIGNATURE OF INVENTOR 201	SIGNATURE OF INVENTOR 202
DATE	DATE	DATE
SIGNATURE OF INVENTOR 203	SIGNATURE OF INVENTOR 204	SIGNATURE OF INVENTOR 205
DATE	DATE	DATE